

Freedom consequentialism : in support of a new measure of utility

A thesis submitted in partial fulfilment of the
requirements for the Degree

of Master of Arts in Philosophy

in the University of Canterbury

Department of Social and Political Sciences

by Daniel McKay

University of Canterbury

2013

Table of contents

Abstract.....	4
Introduction.....	5
Chapter 1: introducing the position.....	10
Chapter 2: moral rights.....	19
Chapter 3: agents.....	27
Chapter 4: other values.....	36
Chapter 5: solutions.....	52
Conclusion.....	66
References.....	67

Acknowledgements

I would like to thank both my supervisors; Carolyn Mason and Michael-John Turp for the help and support they have provided in producing this thesis. They have been an invaluable source of feedback, advice and ideas and I cannot imagine doing this thesis without them.

Abstract

Classical utilitarianism faces significant problems: it ignores moral rights; it cannot take account of all free rational agents; and its focus on happiness means that it dismisses the other things that people value for their own sake. These problems lead to conflicts with autonomy, personal integrity and inconsistencies with the way in which utilitarianism justifies the value of happiness. This thesis seeks to solve these problems by introducing the protection of freedom as a new measure of utility.

Introduction

In this work, I explain and argue in support of a new theory to replace classical utilitarianism. I call this new theory “freedom consequentialism”. Classical utilitarianism has significant shortcomings:

1. Classical utilitarianism ignores moral rights.
2. Classical utilitarianism is not applicable to all actual and potential free rational agents.
3. Classical utilitarianism ignores other things, besides happiness, that people value, which leads to conflicts with personal integrity.

I argue that changing the measure of utility from happiness to freedom solves these shortcomings.

Classical utilitarianism

This work focuses on classical utilitarianism, and the arguments will be directed at that form of utilitarianism, although they apply to some other forms of utilitarianism as well. However, even classical utilitarianism has been developed in different ways by different theorists, which makes it necessary to explain exactly what I mean when I say “classical utilitarianism”.

Classical utilitarianism has many prototypical features (Kagan 17-22; Mill 1879; Scheffler). Classical utilitarianism is a form of evaluative consequentialism; according to any classical utilitarian theory, whether an action is morally right or wrong is dependent on the value of the consequences of that action, as opposed to any other evaluative features, such as whether the action is just, infringes on human rights, or has some other intrinsic quality. Classical utilitarianism is a form of direct consequentialism in that the rightness of the act is dependant on the consequences of that act alone, not the consequences of the agent’s motive or of acting in accordance with a rule recommending that act. Classical utilitarianism judges the total good to come out of an action, not the average per person. It seeks to maximise the good and minimise the bad; the right thing to do is act in the way that produces the highest amount of happiness, after subtracting unhappiness, in the world. Another central tenet of utilitarianism is that benefits to one person matter exactly as much as the same benefits to any other person. It is an agent neutral moral theory; whether something is right or wrong is the same regardless of whose perspective it is looked at from.

Classical utilitarianism is commonly treated as a form of expected value consequentialism. This kind of consequentialism judges whether an action is good, not by reference to its actual consequences, but by reference to its rationally expected consequences. Although the arguments given here are formulated with expected value consequentialism in mind, most of them also apply to actual value consequentialism.

Classical utilitarians generally believe either pleasure and the avoidance of pain, or happiness and the avoidance of unhappiness, to be the only things that have ultimate moral value, and they claim that consequences should be evaluated by them and them alone.¹ While my arguments could be applied as easily to either of these versions of

¹ By “ultimate moral value” I mean non-instrumental value.

utilitarianism, happiness is generally regarded as a more sophisticated, and better, value than pleasure. So, in the interest of not creating a straw man, promoting happiness and reducing unhappiness shall be taken as the measure of utility for classical utilitarianism.

This kind of utilitarianism is generally claimed to be a form of universal consequentialism, meaning that it applies to all persons, or sentient beings, equally (Mill 1879). It does not differentiate on the basis of species, righteousness of character, position of authority, age or intellect. For the purpose of this work, it does not matter whether utilitarianism applies to only persons or more broadly to all sentient beings, so long as it applies, or is intended to apply, at least to all persons.

There are many other forms of utilitarianism, and the arguments used here will not apply to all of them; classical utilitarianism is the primary focus of my critical arguments. Classical utilitarianism still has a great deal of impact on the moral and political landscape, even though philosophers have developed more complicated forms of utilitarianism. Classical utilitarianism is still used as a starting point for most philosophers seeking to improve on some aspect of utilitarianism. Judith Decew, for example, takes classical utilitarianism to be the starting point when arguing for rule utilitarianism (110). Similarly, in *A Theory of Justice*, John Rawls acknowledges that there are many forms of utilitarianism, but contrasts his theory with classical utilitarianism (19). Classical utilitarianism is also one of the simplest forms of utilitarianism. When dealing with something as foundational as what it is that has value, it is best to avoid any unnecessary complications.

Against classical utilitarianism

While there are many problems with, and arguments against, classical utilitarianism, only three are considered in this work: whether utilitarianism can account for rights; the applicability of the theory to all rational agents; and, whether utilitarianism's focus on happiness as the only non-instrumental good is damaging to our integrity. These problems were chosen for several reasons.

Utilitarianism's inability to take account of rights was chosen partially because it is such a commonly raised issue (Brandt 1984; Frey; Griffin). Claims about rights are very widespread in the Western world at the moment, and many of the thought experiments that are used against classical utilitarianism are, at least partially, based on claims about rights. This potential criticism of classical utilitarianism is particularly relevant to my arguments in support of freedom consequentialism, because it is an issue with the theory of value used by utilitarianism. Rights are treated as safeguards of our personhood (Griffin 149). Utilitarianism often ignores the person and focuses only on the happiness they could potentially experience. When this happens, we are reduced, in a sense, to happiness vessels, while our personhood is ignored (Frey 8). While utilitarianism can indeed take account of rights and not all versions ignore them, it can only take account of them in a utilitarian way. Classical utilitarianism often treats rights as having merely instrumental value or as safeguards that can be overridden when there is enough of a utility gain to be made. While this takes account of rights, it misses the point and reduces rights to mere "appendage rights" that are not given the strength that rights theorists think they deserve (Frey 66-67).

Classical utilitarianism's applicability to all free rational agents was the most obvious issue to include. If one wants a moral theory to be universally applicable to all persons, and it seems that utilitarians do, then it must apply equally to all free rational agents in equivalent circumstances.

I use the phrase "free rational agents" in a specific way. When I write that an agent is "free", I mean that the agent has the ability to make choices that are not wholly determined by anything external to that agent and that the agent could have made a different choice in the same circumstances. The term "rational" is potentially misleading. This work does not require that an agent never acts irrationally in order to be a rational agent, only that they have the capacity to think and act rationally. Many types of rationality are important in philosophy, but many philosophers accept some variant of the idea that rationality consists of being able to consider and respond appropriately to reasons (Broome 349; Raz 1985 355). That is what I mean when I use "rational" in this work. As for "agent", the word is used here to mean any self-aware being, that is, a being that possesses consciousness and is able to reflect on that consciousness. This does not preclude non-human animals from being free rational agents. However, the specifics of which non-human animals it would include are best left to those with more empirical knowledge of the subject.

If moral theories in general, and classical utilitarianism in particular, are to apply to all free rational agents, the best way of accomplishing this is to base them on qualities that exist in all free rational agents, not qualities that only exist in humans. In order for a moral theory to apply to all free rational agents, it must not be based solely on human, emotional concerns. To see this, imagine an alien who feels no emotion and yet is a free rational agent.² Classical utilitarianism would have us treat this agent as though it was morally irrelevant, because it does not feel emotions and, thus, cannot be happy or unhappy. What is more, according to the classical utilitarian, morality would have no role on this being's planet and a mass murderer's actions, so long as nobody off-planet hears about them, would be morally neutral.

The last criticism of classical utilitarianism examined in this thesis, considers whether valuing happiness to the exclusion of everything else that people may value is inconsistent with the way the value of happiness is justified. I also consider the claim that ignoring the other things, besides happiness, that people value leads to a conflict between utilitarianism and personal integrity. This criticism was chosen because it is a criticism specifically about the measure of utility used by utilitarianism.

Freedom consequentialism

I introduce a new way to measure utility that is not susceptible to the three objections described above. Any theory developed in response to these objections must give an account of value that is applicable to all free rational agents, safeguards that which is at the heart of personhood, and must not be destructive to people's personal integrity. Freedom consequentialism is based upon something that is common to all free rational agents and at the heart of personhood and integrity: freedom. Free will, as it exists in all free rational agents, is defined in the first chapter. When I refer to

² This thought experiment is discussed in more detail in Chapter 3.

freedom as a value, I mean the exercise of one's free will, rather than the possession of the faculty of free will. To illustrate the difference, imagine a person who has been imprisoned. While this does not affect their faculty of free will, it does affect their ability to exercise that free will, and this is what I am referring to when I talk about freedom.

I argue that freedom has ultimate moral value; an act is bad if it violates a person's freedom and good if it protects a person's freedom or reduces violations of a person's freedom. All other acts are morally neutral. Notice that I am not asserting that what makes an action good is that it increases freedom. Rather, I assert that what makes an action good is that it protects or reduces violations of the freedom a person already has, that is, the freedoms over those things that already belong to them. As a simple example, imagine a man without a great deal of money. He has enough to live on and provide food, shelter, healthcare and other necessities for himself and his family, so there is no concern that this man's life might be taken from him by his circumstances, but he cannot afford luxuries. You, in contrast, have sufficient money to buy any luxuries you desire. Under freedom consequentialism you are not morally required to give this man your own money. While it would be nice of you to give this man some money, you are under no moral obligation to do so as long as his freedom is not impeded by starvation or other health issues brought on by poverty. This position and the reasons for it will be explained fully in the first chapter.

Freedom consequentialism is, as the name suggests, a consequentialist theory. Freedom consequentialism determines whether an act is good or bad based on whether the act protects the freedom of persons or violates it *on balance*. Sometimes it will be a good thing to violate some amount of freedom in order to protect or restore a greater amount of freedom. This greater amount of freedom can be freedom over the same type of thing, such as killing one person to prevent the deaths of three. Alternatively it can be sacrificing freedom over one thing in order to protect freedom over another, more important, thing, such as tearing open a car with the Jaws of Life in order to save a person trapped inside, violating their freedom decide what happens to their property by destroying it in order to preserve their freedom to live. Sometimes these calculations will not be easy, especially when they include freedom over different things. For example, it would not be intuitively easy to determine how many rapes, a serious violation of a person's freedom over their body, are equal to one murder, a violation of a person's freedom to live. However, that is simply the nature of morality; it is hard. Most consequentialist theories face this same problem and while difficult, it is not insurmountable.

Freedom consequentialism would employ something other than a maximising approach. Utilitarianism's maximising approach leaves it vulnerable to a number of criticisms. One of the most serious problems with the maximising approach is that it leads to the demandingness objection, which I discuss later in this work.³

³ The demandingness objection is not the only problem maximising leads to, Samuel Scheffler describes others in his book, *The Rejection of Consequentialism*. However, this issue is not central to the issues under discussion here.

Chapter outline

In this thesis, I examine three major problems with classical utilitarianism and offer a solution in the form of a new kind of consequentialism, which I call “freedom consequentialism”.

The first chapter introduces freedom consequentialism, explains what is meant by “freedom” and explains the kind of freedom that is the foundation for freedom consequentialism. In this chapter, I explain why free rational agents are worthy of moral consideration and why the exercise of their free will has moral worth.

The second chapter discusses the issue of moral rights and whether utilitarianism can adequately take account of them. In this chapter, I examine why rights matter, the conflict between rights and utilitarianism and how utilitarianism tries to take account of rights.

The third chapter looks at whether utilitarianism can be applied to all free rational agents, possible as well as actual. In this chapter, I use a thought experiment to demonstrate the problems utilitarianism has with taking account of moral agents who do not experience happiness or unhappiness.

The fourth chapter looks at whether utilitarianism’s focus on happiness as ultimately valuable, to the exclusion of all else, leads to a conflict with people’s basic integrity and an inconsistency with the way utilitarianism justifies using happiness as the measure of utility. In this chapter, I examine what utilitarianism is committed to saying about cases where someone values something other than happiness, such as art or the pursuit of knowledge. I argue that the utilitarian response is both hypocritical, given the way in which utilitarians justify assigning happiness ultimate value, and damaging to our personal integrity.

The fifth chapter offers a solution to the problems posed in chapters two, three and four and argues for a change from the classical utilitarian theory of value to one based on the protection of freedom. In this chapter, I outline the way in which freedom consequentialism solves the problems presented in this work and discuss potential criticisms that could be raised against it.

Chapter 1: Introducing the position

“Freedom consequentialism” is a new form of utilitarianism based on the protection of the freedom of free rational agents. In this chapter, I explain what is meant by freedom, what kind of freedom ought to be protected, why freedom matters, and, through the use of several examples, I illustrate how freedom consequentialism differs from classical utilitarianism.

Defining terms

Morality

To explain why freedom is morally important, we must first talk about what we mean by morality. I use “morality” to refer to a system of universally applicable rules, maxims or duties that describe the way all moral agents ought to behave. This definition means that morality will not concern itself with things that only some types of moral agents, such as humans, care about, but others may not care about. This does not mean that things that we, as humans, care about, such as fidelity, love, family and many others, do not, or should not, matter to us. These things may be very valuable to us as humans, but it is value of a different kind to moral value. Sometimes these human values, such as loyalty to one’s family or country, may be mistakenly thought of as moral values, but under this definition of morality they are not. This is not to say that we do not care deeply about them, but they are of value to us as humans rather than of value to us as moral agents. It is easy to see how one could mistake these things for moral issues, as our notion of morality seems to have evolved for much the same reason as these other concerns. But, if we want morality to be universal, then moral values must be thought of separately from these specifically human values.⁴ While we ought, first and foremost, to do what we are morally obliged to do, there is no reason why we cannot also care about things that we specifically value as humans.

I assume that all moral agents are worthy of moral consideration. While this is not a controversial claim, it is worth discussing. There are many arguments one could make for why free rational agents deserve moral consideration, but the best reason for assuming this when arguing against classical utilitarianism is that classical utilitarianism itself seems to assume this. When John Stuart Mill argues that happiness is morally valuable, he bases his argument on the fact that we all desire happiness as our ultimate end (1879 49-50). If something matters morally because we all want it, then what we all want must matter. If what we all want matters morally, and matters only by virtue of the fact that we all want it, then what we want is determining what matters morally. So, in a very real sense, we, as moral agents, determine what matters morally. If we determine what matters morally, then it seems that this moral considerability is in some way dependant on us. If we can impart moral considerability in this way, it is reasonable to assume that we, ourselves, matter morally as well. If we are to treat like things alike, then if we matter morally, we must assume that other agents that are like us must matter morally as well. The question then becomes: like us in what way? Utilitarians might be tempted to say they must be like us in the sense that they experience happiness and unhappiness, but happiness, by

⁴ By “human values” I do not mean that no other free rational agent could hold these things to be valuable, only that they are not applicable to all free rational agents that could potentially exist.

the utilitarians own admission, only matters because we all want it. If we are the reason happiness has moral value, the relevant fact about us that makes us matter morally cannot be our ability to experience happiness and unhappiness, it must be something else.

Mill explicitly stated of utilitarianism that it applies to “all of mankind; and not to them only, but, so far as the nature of things admits, to the whole sentient creation” (Mill 1879 17). Assuming that by “sentient” Mill meant capable of conscious experience, this certainly seems like an endorsement of the idea that morality should apply to and take account of all free rational agents. Of course, many utilitarians, including Mill, do not concern themselves with only free rational agents; they concern themselves with other animals as well. This work will not be arguing against giving moral consideration to creatures that are not free rational agents, but neither will it be arguing for it. Rather it will leave the issue of whether other animals are morally valuable or not to one side and focus just on free rational agents. So long as utilitarians think that all free rational agents ought to have moral consideration, and it seems that they do, that is enough for this work to assume this as well.

Freedom

“Freedom” is defined here as the exercise of one’s free will, the making of one’s own choices. Free will has already been defined in the introduction, but freedom is slightly different.⁵ Having free will simply means that one has the mental capacity to make choices for oneself. Freedom, as used here, is being able to use that faculty to make your own choices. To illustrate the difference, consider the case of a man being imprisoned. Presumably being imprisoned does not affect the man’s mental faculties, so his free will is not affected. However, his ability to exercise that free will in the making of his own choices, his freedom, is affected.

In this section, I argue that freedom is central to our agency and our morality. This is related, but slightly different, to free will being central to our agency and morality. It is in the making of choices that our agency is realised, therefore our ability to make choices, our freedom, is central to our agency.

This work will argue that freedom ought to replace happiness as that which has absolute value in the utilitarian model of morality. But, before it can argue that, we must define what kind of freedom ought to matter. Mill said of freedom, “The only freedom which deserves its name, is that of pursuing our own good in our own way, so long as we do not attempt to deprive others of theirs, or impede their efforts to obtain it” (1910 16).⁶

⁵ I assume the truth of either a compatibilist or libertarian account of free will. If these are wrong, and determinism is correct, then it would pose a significant problem not just for this theory, but for almost all realist normative ethical theories.

⁶ It should be mentioned that Mill thought of harm as including emotional suffering caused by the actions of others. This work will not treat emotional hurts in and of themselves as a harm, unless they result from force, such as drugging you. This is because under freedom consequentialism harming someone can be interpreted as violating their freedom and, unless you alter someone’s emotional state by directly interfering with their brain or mind, you have not violated their freedom. Also, the inclusion of emotional harm can lead to serious problems with the Harm Principle, such as whether other people being offended or disturbed by your decision should be considered harm (Holtug 363-364). This means that while it would be wrong to chemically, surgically, electrically, or otherwise directly alter

The problem with Mill's definition is that what it means to deprive others of their freedom is vague. It could be interpreted so broadly as to include the case of a man buying a piece of land and, so, depriving others of their freedom to buy that land. This would be far divorced from what people actually mean when they talk about freedom. Freedom is not the same as getting what you want, and the definition we use should reflect that.

For these kinds of issues not to arise, we must specify what kinds of freedom are to be protected. It is important not to excessively limit the kinds of freedom that ought to be protected. We do not want to define "freedom" so narrowly as to divorce it completely from how the word is actually used. If we were to do that, then we would not really be talking about freedom any more but something else entirely. "Freedom", as it is actually used, involves an ability to pursue many and varied projects, and the definition we come up with must incorporate this if it is to remain true to its common meaning.

This work will limit its concern with freedom to the freedom to do what you will with those things that belong to you: your body; your mind; and, your possessions. As to what choices belong to which people, Mill talked about this when formulating the harm principle. Mill wrote, "In the part which merely concerns himself, his independence is, of right, absolute. Over himself, over his own body and mind, the individual is sovereign" (1910 13). When he said this, he was talking about when the state should and should not intervene, but this could also be used as a guide to what kinds of choices belong to a person. Choices about one's own body and mind certainly seem to be the kind of choices that do belong to a person, although that is still a bit vague, so let us try to flesh it out.⁷

"Sovereignty over one's own body" seems simple enough; this shall be taken to mean that everyone possessing the appropriate faculties to make the decision should be able to have complete freedom over what they do with their body. This includes: whether they live or die; whether they want to keep their organs and body parts; whether, with whom, and in what way, they consent to physically interact; whether they would like to be in pain; whether they would like to move and, to an extent, where they would like to move to; whether they would like to damage their body parts; and whether they would like to perform some action their body is capable of performing, such as speaking or giving birth.⁸

someone's emotions, it would not be wrong to insult them or call them names. In the one case you are taking someone's freedom away and making them feel something, whereas in the other case you are simply saying something, and the emotions they feel have come from their own web of beliefs and attitudes. Of course, because this is a consequentialist theory, if your actions were to lead to consequences that would have an effect on someone's freedom, such as them developing a mental illness, then they could be considered wrong. The point here is not that it is never wrong to do something emotionally hurtful to someone; the point is that it is not intrinsically wrong. This issue is discussed further later in the chapter.

⁷ Philosophers such as Robert Nozick and John Locke base their theories of property on the idea of self-ownership, Locke saying of this, "every man has 'property' in his own 'person'" (Locke 12; Nozick 1974).

⁸ Assuming of course it does not harm the baby.

“Freedom over one’s own mind” includes freedom over what they think about, the continuation, or not, of their consciousness; and what to believe and desire, inasmuch as they can choose these things at all.

Mill did not mention property, but it certainly seems to be something one has dominion over. If people cannot own property, then obviously it does not belong in that domain, and it would not be wrong to take or destroy some object because it has no owner. However, assuming that people can own property, it is by definition theirs and belongs in the domain of things they ought to have freedom over.⁹ This would include: whether or not they would like to have it damaged or destroyed; who they would like to use it; and whether they would like to sell, trade, or give it away, to whom and for what price.

Part of one’s body could also potentially become property if removed and sold, those parts could then transfer from being another person’s property into being part of their body in cases of transplant. This is because organs shall be assumed not to have a special moral status. Unless there is some good reason why an organ, once removed from the body, is morally different from any other object, they ought to be treated similarly. Granted they can be treated as important objects if they can save lives, but not as morally special for no other reason than that they were once a part of a human body. There does not seem to be a good reason to believe that organs are morally different from other objects, so they shall be assumed not to be.

While you may give away your property, those things that are externally yours, you cannot give away that which is internally yours, your freedom (Kant 1996 38). It is impossible for one to become a slave, even voluntarily, as it is impossible for one person to own another. Even if you were to sign away your freedom, it would still be yours. The choices to move, speak, live, along with many others would still be yours to make and no one else’s. No one can take these choices from you, they are always yours to make. This is different from selling organs, however, because, assuming the donor consents, once removed from the body, the organ is just an object. Freedom on the other hand is not just an object, it is part of our essence as agents, and not something one can barter or trade away.

Before moving on to the next section there is something that ought to be clarified. In this section it has been mentioned that the goal of freedom consequentialism is to protect freedom rather than to increase it. Protecting freedom rather than maximising it may seem odd given the utilitarian starting point of this thesis. The reason “protect” has been used instead of “maximise” is simple; the type of freedom at issue is the freedom for everyone to do what they will with those things that already belong to them. Trying to increase this kind of freedom amounts to the same thing as protecting it, because it is freedom over those things that *already belong to agents*. Freedom consequentialism will indeed support increasing the freedom of everyone, but it will only support increasing a specific type of freedom, freedom over what they already rightfully own.¹⁰

⁹ Immanuel Kant (1996 47), Nozick (1974) and many other philosophers believe that we can own property, but persuasive objections have been raised against these accounts of how we can come to own property. I have not provided an argument for how we can come to own property, because I am not sure there is such an argument.

¹⁰ As opposed to things obtained by theft or fraud.

Why freedom matters

This work primarily argues for freedom having ultimate value instead of happiness by presenting problems for happiness-based utilitarianism that freedom consequentialism can solve. However, other potential solutions to the problems faced by classical utilitarianism are presented in this work, and while some of these potential solutions maintain happiness as that which has ultimate value, some do not. However, there are reasons to value freedom beyond simply that it solves some problems with classical utilitarianism. This section will argue one compelling reason to value freedom is that it is central to our personhood and to morality.

Like many moral theories, including deontology (Kant 1997 2-3) and utilitarianism (Mill 1897 17), this work assumes that the correct moral theory will be universal. If a moral theory is to be universal, then the theory must apply to all moral agents, both those we know to exist and those that may possibly exist. The best way for a moral theory to take account of all moral agents, is for it to be based on something that is present in all moral agents. This limits the options for what can be of moral value dramatically. Later, I argue that happiness is not applicable to all moral agents, but in the meantime let us focus on what we know is common to all moral agents.

If moral agents are, by definition, free and rational, then we already have two things that are common to all moral agents. To be a rational agent one must have preferences.¹¹ A being who lacked preferences, and so only performed actions at random, or who performed no actions at all would not be a free rational agent. The idea of agency is tied up in acting, and the idea of being rational is tied up in having reasons for your actions. It could certainly be argued that in order for agents to have reasons, or at least motivating reasons, they must have preferences. So, for the sake of argument, we will say that without preferences one cannot be a rational agent. There is already a normative ethical theory based on preferences: preference utilitarianism. Preference utilitarianism uses preferences as a guide for what is morally right, and it seems to be applicable to all free rational agents and thus all moral agents (Singer 1993 94). Since preference utilitarianism already provides a theory of morality that applies to all free rational agents, freedom consequentialism must offer more than freedom's commonality to all free rational agents to offer a worthwhile alternative theory of morality. In the next section, I argue that freedom is a better candidate for ultimate moral value than preferences because of its importance to morality.

What sets freedom apart from other characteristics common to moral agents, such as having preferences, is its place at the heart of what it is to be moral. It is impossible to think of an agent being morally praiseworthy or blameworthy for actions they have no choice over. This may sound like a pointless reiteration of the idea that only free rational agents are moral agents, but it is not. To illustrate how freedom is central to morality, let us imagine the case of an artificial intelligence that has free will in all cases but one. It has been programmed, without its knowledge, to kill a small child upon seeing it. When it kills the child, assuming there really was no way that it could choose to do otherwise, and that it did not know about this aspect of its programming

¹¹ One could argue that someone could be a free rational agent without having preferences. Whether this is possible depends largely on what one means by preferences, but, for this work, I am conceding that preference utilitarianism may well be based on something universal to moral agents and, instead, arguing that the protection of freedom has other advantages over preference utilitarianism.

beforehand, it cannot be said to have acted wrongly. There can be said to be bad consequences, but a volcanic eruption can have bad consequences, and we can no more blame the artificial intelligence for killing the child than we can blame the volcano for erupting. “Ought implies can”, after all, and if the entity in question could not have done otherwise, then we cannot say that he ought to have done otherwise.

The notion of “ought implies can” is a simple concept that is not only about freedom, but also about ability. What is often left out is that ‘ought’ also implies ‘can do otherwise’. Just as it makes no sense to talk about one doing wrong when they had no choice in the matter, it makes no sense to talk of one doing right when they could not choose to do wrong. To illustrate this, imagine a being, which we will call an ‘angel’, which must always do good and never evil. The angel always helps those who need it and acts in the best interest of others. Hurting or stealing from people would never even occur to it, it is simply not a part of its nature. Even if someone were to suggest the idea to the angel, it could no more do evil than we could flap our arms and fly. But, is the angel acting morally? It might seem tempting to say that it is and we might well want to praise the angel for its actions. But, it would amount to the same as praising the rain for falling or the sun for shining; it might please us to do so, but they did not get any choice in the matter. It does not make sense to say that someone can be acting morally, or can be praiseworthy, when they could not have done anything else (Kant 2010 323).

Freedom’s central role in morality is not the only reason to accept freedom consequentialism over preference utilitarianism. The objections raised in chapter two and four apply to both classical utilitarianism and preference utilitarianism, and the issue of why to accept freedom consequentialism over preference utilitarianism is discussed further in chapter five.

Examples

The position that this work champions is a complicated one, and there are a few things that should be spelt out to make it easier to understand what is being argued for. The easiest way to explain the differences between freedom consequentialism and classical utilitarianism is to look at how each one would recommend people act in several different sets of circumstances. This section looks at just that, using some of the most popular cases used to demonstrate, or criticise, classical utilitarianism.

Trolley problem

The first case examined is the classic trolley problem. A trolley car is careening out of control. On this trolley there are five people. The tracks that the trolley is on lead off a cliff and the trolley cannot stop in time to prevent it from going off the cliff. If you do nothing, all five people will definitely die. However, there is a lever you can pull that will divert the trolley onto another set of tracks and save all five people on board. On the second set of tracks there is a hiker. Sheer rock walls surround the tracks, so the hiker cannot escape. If you pull the lever, the hiker will die. So, either you pull the lever and save the five people on the trolley but kill the hiker, or, you do nothing and let the trolley go over the cliff and the five people on board will die.

Classical utilitarianism recommends that you pull the lever. Assuming all of the people on the tracks are of roughly equal moral worth, you can conclude that the life of the five people easily outweighs the life of the one. The calculations here are simple. The one person has less happiness in their life and less family and friends to mourn them than the five, so a world in which five people are saved is better than one in which one person is saved. Because classical utilitarianism draws no distinction between action and inaction the choice is simply between one person living and five people living.

Freedom consequentialism will come to almost the same conclusion, but for slightly different reasons. As with classical utilitarianism, it is a form of consequentialism and it does not draw a distinction between action and inaction. The only difference is that you would weigh the freedom the action would protect rather than the happiness that would result. In this case, either five people will die against their will, or one person will die against their will. So, freedom consequentialism will also recommend you pull the lever, but the calculations are based on the protection of freedom instead of the maximisation of happiness.

Christians and lions

Next, let us look at a case where classical utilitarianism and freedom consequentialism differ, the case of the Christians and the lions. A huge Roman coliseum is filled with thousands of Romans who want, more than anything else, to see some Christians get ripped apart and eaten by lions. They have the lions all ready and some Christians on hand as well. Assuming that the case has been set up properly, making sure that word of the Christians being eaten does not spread to those whom it would upset, classical utilitarianism must say that it is right to throw the Christians to the lions. If it does not say this, because the happiness lost by the Christians is greater than that gained by the Romans, all one needs to do is increase the number of Romans watching. This is because utilitarianism simply calculates the amount of happiness versus unhappiness that follows from different possible actions and requires agents to perform the action that will produce the most happiness minus unhappiness. It does not care if the unhappiness of some would be very great; it just looks at aggregate happiness versus unhappiness.¹²

Freedom consequentialism takes a different stance on this case. It also looks only at aggregate value, but because what has value is different, it can say different things about cases such as this. First, freedom consequentialism requires that we consider whether any of the Christians' freedoms are being violated. Clearly, they have been deprived of the freedom to choose whether to live or die. Second, freedom consequentialism requires that we consider whether not throwing Christians would violate the Romans' freedom in any way. In this case, it clearly will not. The choice of whether the Christians are torn apart is not the Romans to make; they have no dominion over the lives of the Christians. The Romans certainly have freedom of action, so they may pursue any form of entertainment that does not involve the violation of another's freedom. However, freedom of action does not extend to violating the freedom of another. It is no violation of a person's freedom to prevent

¹² This is not the case for all forms of utilitarianism, but, as mentioned in the introduction, it is the case for the type of utilitarianism at issue here.

them from murdering another person, because whether another person lives or dies was not their choice to make in the first place. So, with significant violation of freedom on one side, and no violation of freedom on the other, freedom consequentialism entails that throwing the Christians to the lions is wrong. This is because only a certain type of freedom matters, the freedom over those things that belong to oneself.

Infidelity

The next example is not a thought experiment that only comes up in philosophy classrooms; it is a moral issue that comes up often in real life. It is the case of having an affair. Carrying on an affair without your partner knowing is an interesting case to look at, because it contains two separate elements for moral scrutiny: the element of betraying your partner's trust in you, and the element of lying about it. Most moral theories condemn lying as wrong, and few favour of cheating on your partner. Utilitarianism differs from many moral theories as it recommends lying whenever lying has the best consequences. Utilitarianism even condones cheating in circumstances where you do not feel guilty, no one finds out about it, and the people you cheat with are not hurt by being involved in your indiscretion. However, utilitarianism will not recommend cheating all the time, or even often. When someone cheats on a partner it usually causes a great deal of emotional hurt to someone. In a case-by-case analysis, utilitarianism would generally say that cheating on your partner is wrong because someone is likely to get hurt. However, if you did cheat on your partner, utilitarianism may recommend lying about it to protect their feelings and save yourself the consequences of them finding out, depending on whether they are likely to find out from another source. Because freedom consequentialism is concerned with the violation of freedom over what belongs to you, not happiness, and because you have no claim over what your partner does so long as it does not affect your freedom, freedom consequentialism takes the position that, in and of itself, cheating on your partner is not a moral issue.

Let us imagine a man, who we will call Peter, who has a girlfriend who we will call Amy. The two of them are happy together and are very much in love. However, Peter is a cheater. He likes to cheat on Amy with his mistress, who we will call Beth. He has told Amy about this, and she was distraught, she loves Peter very much and she wants him to be faithful to her. Amy stays with Peter even though she knows he is cheating on her because she loves him and does not want to lose him. Peter feels guilty about cheating on Amy, who he loves, but he does it anyway. Beth feels guilty about her part in all this but she tells herself that she is not the one cheating and that if it was not her it would be someone else. Now what are we to say about the moral implications of this whole mess? According to freedom consequentialism, absolutely nothing. Nobody has had their freedom over those things that belong to them violated. Peter is exercising his freedom to decide who he wants to have sex with, as is Amy, and as is Beth. Amy has no claim over what Peter does with his body. He does not, and indeed cannot, belong to her. He is his own man and has sole dominion over whom he chooses to sleep with.

Earlier we looked at the difference between moral issues, those that apply to all free rational agents, and things that we, as humans, care about. Under freedom consequentialism, cheating falls under the second category. We humans usually care

about whether our partners sleep with other people, but this is not necessarily the case for all free rational agents. It is easy to conceive of free rational agents that do not form strong emotional ties with their romantic partners, or indeed who reproduce without the need for a partner at all. So, while fidelity may matter to us on a personal level, it is not a moral issue.

It is important to stress that in the example of Peter and Amy there is an equal power relationship between both partners. There are certainly many relationships in the world where this is not the case and depending on the nature of the relationship one partner may not be truly free to leave the other, but that is not the type of case we are looking at here. Lying does not occur in this example because lying to people raises issues of autonomy and how people make decisions and would unnecessarily complicate the example. It is also assumed that there are no sexually transmitted diseases involved, and that there will not be some other terrible consequence of Peter's affair. This is because this is a consequentialist theory and any action can be wrong if it has bad consequences. So, while cheating may not be wrong in and of itself, if Peter contracted a sexually transmitted disease, then it could well be wrong because there could be significant violations to the freedom of both himself and his partner if he then transmits that disease to her.

It should be mentioned that freedom consequentialism does not suggest that it is a good thing for Peter to have an affair, or that we cannot dislike him for having one. It says that it is simply not a moral issue. We could still think Peter is unpleasant, still hate him, still not want to associate with him, still tell his girlfriend to leave him, or even name and shame him on the internet. But, morality would have no say in the matter, and we would be wrong to force him to be a nicer person. People have freedom over what they say and how they act, so long as it does not violate the freedom of another. This means that people have the right to be vindictive, nasty, petty, and cruel so long as they do not violate anyone's freedom in doing so. These are things we may not like as human beings, but, as discussed above, that is not the same as them being wrong.

Conclusion

Freedom is important because of its central role to both our agency and to morality. Freedom consequentialism is a consequentialist theory that uses the protection of your freedom over things that belong to you, rather than the maximisation of happiness, as the measure of utility. Because of its different measure of utility, it can recommend different actions to classical utilitarianism. Now that freedom consequentialism has been introduced we can move on to looking at some of the problems it can solve.

Chapter 2: Moral rights

Classical utilitarianism conflicts with widespread intuitions about moral rights. Rights conflict with utilitarianism when an agent has a right to act in a way that fails to maximise happiness or when the course of action recommended by utilitarianism would violate a person's rights. Utilitarians do not take these conflicts to show that there is a problem with utilitarianism. Some utilitarians respond by rejecting the claim that there is a useful, consequence independent, concept of rights. Others suggest altering classical utilitarianism so that it becomes compatible with the existence of rights, by developing a two-level account of utilitarianism or by trying to work rights into a utilitarian theory. However, none of these responses succeeds, leaving classical utilitarianism with a problem it cannot resolve.

Defining rights

Before we can look at the conflict between moral rights and classical utilitarianism, we must first define what a right is. The term "moral rights" will be used rather than "human rights", because the term "human rights" seems to assume that other potential persons do not have the same innate moral rights that humans do, and this work will not be making any such assumption.

Under the classical conception of rights, to have a right is to have an entitlement to perform certain actions or be in certain states and to confer upon everyone, including yourself, a duty to allow you to perform certain actions or be in certain states (Sumner 23). For example, if you have a right to live, that confers on everyone, including yourself, a duty not to kill you.

The modern conception of rights is a lot more complicated (Sumner 23; Wellman 98). According to it, rights are systems of duties, entitlements and choices. Under the modern conception of rights, a right to live confers on others a duty not to kill you but also entails that you can make a choice to die. This modern conception of rights seems to be less likely to conflict with utilitarianism. To illustrate this, imagine a man on his deathbed in terrible pain who wants desperately to die. Utilitarianism clearly favours letting, and perhaps even helping, the man kill himself, but the classical conception of rights would say that it would be a violation of his right to live for him to kill himself; thus, it would be immoral for him to act in this way.¹³ The modern conception of rights used here, however, would allow him to take his own life as long as he made an autonomous decision to do so. This work assumes the truth of this modern conception of rights precisely because it conflicts less with utilitarianism, and using the classical conception would make a straw man of the position that utilitarianism and moral rights are in significant conflict.

What are rights and why do they matter?

Opinion is divided over what qualifies as a right. People claim all sorts of things as rights (Sumner 22). Nudists might make a claim to a right to freedom of expression to justify their claim that they should be allowed to run about naked, while their opponents might claim that public nudity violates their right to live in a decent

¹³ This is assuming that the man killing himself will not have significant consequences for other people.

society. For almost every social issue, people on either side of it may claim that one of their rights is being violated. Because of this tendency to call practically anything a right, there needs to be some way of determining what makes something a right or not a right.

In order to determine what qualifies as a right, we need to determine what makes rights valuable in the first place. Rights theorists justify the value of rights by saying that they are the safeguards of something of fundamental importance to our personhood. Some theorists think that what rights safeguard is our autonomy (Sumner 23). Other rights theorists, such as James Griffin, describe what gives rights their value, and what it is that they safeguard, as our freedom (61).

Definitions of freedom and of autonomy conflict; while something might be a safeguard to one person's definition of autonomy it might not be to another's. So, in order to determine what qualifies as a right, as well as precisely what it is that gives rights their value in the first place, we must first define autonomy and freedom.

Philosophers define personal autonomy in different ways, but these definitions have several things in common. They all seek to limit which of an individual's actions can be thought of as autonomous. For example, according to Harry Frankfurt's hierarchical model, an agent acts autonomously if and only if their first order desire to commit the act is sanctioned by a second order volition endorsing that desire (12-25). Whereas Ekstrom claims that an agent acts autonomously if the preferences that lead the agent to action cohere with their character in general (Ekstrom). Some other models seek to limit what agents and actions can be considered autonomous, for example, it has been suggested that agents must have "normative competency", the ability to determine right from wrong, in order to be considered autonomous (Wolf 1990). On each of these models, for an action to be considered autonomous, it is not enough for the action to be of the individual's choosing, it must also be supported by higher order desires, or cohere with their character, or the agent in question must have certain competencies. Under most definitions of personal autonomy, it seems that actions may not be considered autonomous if they are performed at random, on a whim, or, under some definitions of autonomy, if they are in violation of moral law. This does not seem to fit with how rights theorists and laypeople talk about rights. Moral rights are thought to protect an individual's ability to act in a certain way when they act on a whim or even act at random. The only constraint that does seem to apply is that no one has a right to violate the right of another. Even with this constraint, rights protect more than what is traditionally described as "autonomy". So, "autonomy" is not the right word for describing what it is that rights protect and what gives them their value.

James Griffin suggests that it is freedom that makes rights valuable. In order to determine if this is true, we must first determine what he means by freedom. Griffin uses a Kantian definition of freedom, saying that freedom is "independence of constraint by another's choice" and that the freedom of every agent must be able to coexist with the freedom of every other agent under universal law (61). It seems safe to say that when Griffin says that freedom is what makes rights valuable he is talking about a Kantian kind of freedom, so we must define what Immanuel Kant meant by "freedom".

Kant's definition of freedom requires the agent in question to be free from all external influences and that the agent's actions must not be wholly causally determined by what has preceded them or determined by their nature (Kant 1997, 2010; Pereboom 537; Silber 71). According to Kant, to be free, a person must be free from determination by all alien or antecedent factors, such that the choice made is theirs and theirs alone (Silber 72). Also according to Kant, the only free choice is a choice made in accordance with moral law, not one that simply follows our desires (Morrisson 129). Freedom of this sort could potentially be thought of as what gives rights their value, but Kantian considerations of universal law and acting from duty do not fit with the way that most rights theorists, and laypeople, think about rights. Rights protect a person's ability to do, or not do, or have, or not have, a certain thing regardless of what motivated them to do, or not do, or have, or not have, that thing. Also, rights safeguard your freedom to do, or not do, or have, or not have, certain things, even when those things are not in accordance with a Kantian conception of moral law. So what rights protect, and what gives them their value, does not seem to be freedom in the Kantian sense.

Neither "autonomy" as it is traditionally defined, nor "freedom" as defined by Kantians, seems to capture exactly what it is that rights safeguard and what gives them their value. However, they both seem to come close. When Griffin's and Leonard Sumner's explanations of why rights are important, are considered alongside the kinds of things that are usually asserted as rights, it seems that the value of rights is derived from some kind freedom. The ability to act independently from constraint is at the heart of theories about rights. However, this freedom is not an absolute one. Rights theories always include rights over certain things, but not others. Rights theorists agree that one has a right to live, or a right not to be killed, but that one does not have a right to have their name spelled out in twenty-foot letters on the moon (Frey; Griffin; Osiatynski; Sumner; Raz 1985). Rights also do not include the right to violate the rights of another person. So, it seems that what rights protect, and what gives them their value, is some kind of innate freedom over a limited set of things. This is important, as it connects rights with freedom consequentialism, the theory this thesis is proposing.

A list of rights

While we now have a basis for what makes something a right, we still do not have a definitive list of what is and is not a right. Developing a perfect system of rights is a project on its own. So, rather than spend a lot of time analysing each potential right and determining its merit, let us make a list of the most widely agreed upon rights. Those that most rights theorists, as well as most people in general, agree are definitely rights, and use those as examples of apparent conflicts between classical utilitarianism and rights.

The most obvious right that needs to be taken account of is the right to life. This is certainly the most commonly asserted right, and it seems that if there is anything that is an innate right, it is this. Because we are using the more modern conception of rights, the right to life amounts to a choice between living and dying, so the right to life is not incompatible with the right to have an assisted suicide. The other rights that this work focuses on are the right to: safety from violence; free thought; free speech; freedom of action; freedom of religion; healthcare; education; property ownership and

control; expression of one's sexual orientation; adequate housing; freedom of movement; control over one's own body; engage in consensual contracts with other rational beings; and raise children.

This list has been put together from a number of works (Frey; Griffin; Osiatynski; Sumner) as well as the common understanding of moral rights. There may well be other rights that ought to be on the list, and there might be good reasons why some of these things ought not to be considered rights. However, this should not be too much of a problem. If the point is simply to show that utilitarianism comes into conflict with moral rights, then all we need is a list that largely overlaps with what most rights theorists, as well as laypeople, think of as rights. Then, as long as utilitarianism comes into conflict with enough of these supposed rights, it can be considered having a conflict with moral rights even if some of the members of the list are not really rights at all. This does seem to be such a list, so it will be adequate for the purposes of this work.

The conflict between rights and utilitarianism

There are many instances of utilitarianism and rights coming into conflict.¹⁴ One issue where moral rights and utilitarianism conflict is that of smoking. Utilitarians ought to strongly oppose smoking. Smoking causes debilitating and fatal diseases, costs a lot of money, often for those who can ill afford it, and does not have a pleasant high associated with it like some other recreational drugs. Banning smoking might make many people angry in the short term, but a ban would almost certainly have long-term benefits for society from a utilitarian perspective. So why do we not ban smoking? One of the main arguments given in defence of allowing smoking is that banning it would violate our moral rights. Rights to freedom of action and freedom to control one's own body seem to demand that people be allowed to smoke, at least as long as they are not breathing their smoke onto other people. These two philosophies come into conflict often on this issue, whenever a smoking ban or a tax on cigarettes is proposed.

Refusing medical treatment is another issue where moral rights and utilitarianism seem at odds with one another. Refusing blood transfusions, transplants, and other procedures because of religious, or otherwise personal, beliefs is a practice strongly grounded in moral rights. Rights to determine what one does with one's own body and, in some circumstances, rights to religious freedom, demand that people be allowed to refuse treatment. Utilitarianism, on the other hand, opposes this. In cases where the person in question can be cured and will likely go on to lead a long life without any long term ill effects, saving the person's life, even against their wishes, is the utilitarian thing to do. Any temporary unhappiness they may suffer because their rights were violated will be hugely outweighed by the happiness they, and those who care about them, will gain in the long term by them being alive. One might make the argument that if this were general practice, it might outrage many people who hold these kinds of beliefs. If this would indeed outweigh the happiness gained by saving several extra lives, and that is certainly up for debate, it would still only be an

¹⁴ The most obvious conflicts between rights and classical utilitarianism occur in the criminal justice system. This work does not focus on criminal justice, or other forms of punishment, because such conflicts involve justice, and considering issues of justice will unnecessarily complicate the issue of utilitarianism and rights.

argument for not making it public policy; it would not be an argument against one or more doctors giving covert blood transfusions so long as the public did not find out about them. It seems clear that there are cases where utilitarianism would favour saving people's lives whether they want them saved or not, whereas rights-based theories would be staunchly opposed to overriding autonomous decisions in this way.

These conflicts between rights and classical utilitarianism demonstrate that utilitarianism can lead to a paternalistic view of morality; because utilitarians only value maximising happiness, and not autonomy or freedom, utilitarianism will often recommend acting against someone's wishes if the action is in the person's best interests.¹⁵ Moral rights, on the other hand, rest on the assumption that it is morally right for people to have the freedom to make their own decisions regardless of whether their decisions make them and others happiest in the long run. This difference is one of the most important reasons that utilitarianism and rights cannot coexist, and it arises, at least in part, because of the measure of utility used by utilitarians. Utilitarians see no problem in going against people's wishes for their own good, as long as this maximises happiness, because utilitarianism does not give non-instrumental moral consideration to freedom or autonomy, only happiness.

When confronted with this criticism, utilitarians have several options. They could insist that rights are not an important part of a moral theory, try to take account of rights using a 'two-level theory' or they could try to incorporate rights into utilitarianism as things which have some non-instrumental value. These options are discussed below.

The utilitarian's responses

Biting the bullet

The simplest way to respond to the apparent conflict between rights and utilitarianism is to just bite the bullet and say that rights either do not exist or are morally irrelevant. This seems a logical response from the utilitarian. This criticism amounts to an appeal to intuitions after all, so the utilitarian could easily say that this particular intuition is simply wrong; that is, there are no inherent moral rights. This is the starting position of classical utilitarianism, in that it holds that only happiness has ultimate value, and therefore rights do not. But, it does seem that by taking this stance utilitarianism is missing something important. Raymond Frey criticises utilitarianism, saying that it reduces people to mere happiness vessels and ignores much of what is essential to personhood (Frey 8). When utilitarianism ignores rights, which are grounded in, and safeguards of, our freedom, they ignore something fundamental to morality (Griffin 61; Sumner 23).

A preferable approach would be for utilitarianism to take account of rights in some way. They can do this by either using rights as rules for acting because doing so will maximize utility overall, or by assigning rights some amount of non-instrumental value that can be weighed against potential utility gains.

¹⁵ "Paternalism" here refers not only to actions taken by governments or employers, but also to any action where one person makes decisions for other people, without their consent, for those people's own good. Utilitarianism may oppose paternalism certainly, but only if it does not maximise utility.

Rights as a guide

One suggested response to the conflict between rights and utilitarianism is that we should use utilitarianism at a critical level, but we should appeal to rights at the everyday, intuitive, level as part of a system of rules or heuristics (Frey 70). Rule utilitarians give a range of reasons for relying on rules that maximise utility rather than carrying out a utilitarian calculus before every action. These reasons might be used to argue that we should take a rights-based approach rather than using classical utilitarianism to make moral decisions in our everyday lives. Four reasons commonly given by rule utilitarians not to rely on the utilitarian calculus are as follows (Hooker 142-143). First, we often do not have the relevant information to make good utilitarian calculations. Second, we often do not have the time to deliberate and perform utilitarian calculations before making moral decisions. Third, our human biases lead us to make mistakes in our utilitarian calculations. Finally, if it were well known that we made our decisions using a utilitarian calculus, other people may not be inclined to trust us. Although these problems may work as an argument for using some form of rule utilitarianism as a guide for acting, they do not provide a good argument for a rights-based theory of morality being a part of that rule utilitarian framework. It is highly counterintuitive to claim that the best way to achieve a utilitarian end is to make decisions based on a system that is diametrically opposed to utilitarianism. Even if utilitarianism were not a good way of making moral decisions, and a system of heuristics were required, it seems unlikely that moral rights would feature in such a system, as they are completely opposed to the consequentialist aims of utilitarianism.

Richard Brandt argues in support of a different theory about the relationship between utilitarianism and moral rights. Brandt posits a form of rule utilitarianism where whether an act is right or wrong is based on whether it is in line with the best possible moral code, where 'best' means that which produces the most utility (Brandt 1984). He believes that this type of utilitarianism could take account of moral rights. Brandt argues that the way we ought to do determine the correct place of rights in morality is to determine whether the ideal moral code would say we ought to use moral rights as a guide to our actions and when, if ever, we should violate them.

Brandt's theory relies on determining whether rules for acting are in line with the ideal moral code, but establishing what the ideal moral code is would be very difficult. It appears as if Brandt is saying he has a new ethical theory, but that theory amounts to "we should come up with the perfect ethical theory". Another problem is that if the ideal moral code determined by whether it produced the most utility, we do not have a good reason to believe that a rule saying "never violate moral rights" would be in line with it. We have already seen that utilitarianism often requires us to violate moral rights when doing so will increase utility, and it does not seem likely that the ideal moral code, from a utilitarian perspective, would no longer require us to violate moral rights to increase utility.

Rights as valuable

Another way in which utilitarians could attempt to take account of rights is by giving rights actual value and arguing that people ought not violate someone else's rights unless there would be sufficiently dire moral consequences from them failing to do so. This would certainly bring utilitarianism more into line with how we actually think about rights in the real world, but it has several problems of its own.

One of the biggest problems with trying to weigh rights and utility against each other is determining when it is acceptable to violate a person's rights to maximise utility. Presumably the amount of potential happiness needed to justify the violation of a right would depend on what right was to be violated. For example, it is surely not as bad to imprison someone for a day as it is to kill them. But, how much happiness would justify the violation of each right? Regardless of the specific amount of happiness required, there is a sense in which thinking of rights as a safeguard that can be overcome by enough happiness belittles their importance. Frey argues that any direct consequentialist account of rights would reduce them to mere appendages to the theory (Frey 66-67). Frey's point is that rights are meant to be, if not absolute, then very strong and robust. They are intended to be a trump that outranks other considerations, because they are there to protect something central to our personhood and, therefore, our moral agency. Rights are not meant to be a hurdle for utilitarianism to jump over or some sort of check on when it is okay for utilitarians to claim that the right thing to do is to maximise happiness.

Even if we could develop an account of utilitarianism that makes the theory consistent with moral rights, there would still be the issue of whether acting is morally distinct from not acting. Consequentialist theories like classical utilitarianism do not distinguish between action and inaction; according to utilitarianism you are just as responsible for a man's death if you see him drowning and do nothing as if you hold him under the water yourself. Rights-based theories are not consequentialist, so they draw a distinction between action and inaction. Because of this disagreement, utilitarianism can never truly take account of rights.

To illustrate how the distinction between action and inaction brings rights and utilitarianism into conflict, let us engage in a thought experiment. Let us imagine a man who has information about a bomb about to explode. The police have found him and brought him in for questioning. The man refuses to tell the police anything because the terrorists who planted the bomb have kidnapped his son to keep him quiet. Someone suggests torturing the man to extract the information. The man is a bit of a pushover. He would tell the police where the bomb is if tortured for even a short time, so if he is tortured, there will be time to prevent the bomb going off. Most rights-based theories would tell us that we must not torture the man, as doing so would be a violation of his rights. Utilitarianism would tell us that we must torture him because the consequences of the bomb going off will be worse than the consequences of torturing him. The issue with trying to combine utilitarianism and rights is that one draws a distinction between action and inaction and one does not, we can see how this causes conflict by applying it to the case of the man with information about the bomb. Utilitarianism does not need to argue that this is a case where the potential utility gain from saving the would-be bomb victims outweighs the right of the man not to be tortured. Because utilitarianism draws no distinction between action

and inaction, it can say that by not torturing him you are violating the rights of all those who would be blown up and that by torturing him, we are choosing the possible world where fewer people's rights are violated. The fact rights-based theories draw a distinction between action and inaction makes it hard for any consequentialist theory to truly take account of rights, but the combination of that and the fact that classical utilitarianism values something so different to what is protected by moral rights, means that utilitarianism cannot coexist with any robust theory of moral rights.

Conclusion

Classical utilitarianism is unable to take account of moral rights. The paternalistic nature of utilitarianism and its failure to distinguish between action and inaction both contribute to its conflicts with moral rights. These factors, combined with utilitarianism's focus on happiness, to the extent that it treats people as mere happiness vessels, mean that utilitarianism is completely incompatible with any kind of robust moral rights. Utilitarians may respond that this is not to their disadvantage, but the kind of freedom moral rights are grounded in is important to us as moral agents, and, in not taking account of them in some way, the utilitarian treats people as merely happiness vessels. The utilitarian has several responses to this objection, but none of them seem to work. The responses relegate rights to a hurdle to be overcome or introduce some kind of dubious second-order system of morality. Objecting to utilitarianism because of the way in which it conflicts with rights relies on an appeal to a moral intuition, and that intuition could well be wrong, but it appeals to a very strong intuition, and one that is connected to something deeply important to our status as moral creatures. The issue of taking account of rights in a consequentialist framework will be discussed further in chapter 5. In the next chapter I examine whether classical utilitarianism can take account of all potential moral agents.

Chapter 3: Agents

Free rational agents that cannot experience happiness or unhappiness pose a significant problem for classical utilitarianism. Morality, if it is to be universal, must apply to, and take account of, all free rational agents. In this chapter, I show that classical utilitarianism, because it values happiness as the only non-instrumental good, cannot do this. After defining relevant terms, I consider whether utilitarianism is actually intended to take account of all free rational agents. Next, I use a thought experiment about free rational agents that cannot experience happiness or unhappiness to demonstrate the problems classical utilitarianism has with applying to all free rational agents. I examine the responses utilitarians can make to this objection, and I show that they do not effectively deal with the issues raised by the thought experiment. These responses include: claiming that an agent cannot be free and rational without being able to experience happiness; redefining happiness; questioning whether we can imagine such beings; and biting the bullet.

Defining terms

This work talks about normative theories being “able to take account of” and “applying to” free rational agents. A theory will be considered to “apply to” all free rational agents if all free rational agents can act morally under that theory, and are able to follow the dictates of that moral theory; that is, a moral theory applies to all free rational agents if and only if all free rational agents are moral agents. A theory will be considered to “take account of” all moral agents if that theory ascribes moral worth to all free rational agents.

The arguments in this chapter raise objections to classical utilitarianism only, not preference utilitarianism. Preference utilitarianism seems to be able to take account of and apply to all free rational agents. Ignoring the utilitarian’s option to switch to preference utilitarianism may seem like making a straw man of the utilitarian’s position, but it is not. This chapter demonstrates the problem classical utilitarianism has with taking account of all free rational agents, and it goes on to show that this problem is fixable by changing what has value from happiness to something else. Arguments in other chapters introduce objections to preference utilitarianism, but this chapter focusses only on classical utilitarianism.

Why applying universally matters

Before considering whether classical utilitarianism can apply to and take account of all free rational agents, we should first discuss whether it was ever intended to. If classical utilitarians do not claim that their theory applies to and takes account of all free rational agents, then the thought experiment presented here is much less of a problem for them than if they do.

John Stuart Mill talks about this very issue; in *Utilitarianism* he claims that the theory applies to “all of mankind; and not to them only, but, so far as the nature of things admits, to the whole sentient creation.” (Mill 1879 17). This certainly looks like an endorsement of the claim that utilitarianism is universal and applies to at least all free rational agents and possibly more things besides, depending on his understanding of

the word “sentient”.¹⁶ The issue here is the phrase “so far as the nature of things admits”. This could be taken to mean either “as far as possible”, or, “so long as they also desire happiness and the avoidance of unhappiness as the ultimate end”. Putting aside the debate about whether we have happiness and the avoidance of unhappiness as our ultimate end, it seems that this second reading is unlikely. It seems much more likely that Mill meant simply, *as far as possible*. This interpretation of Mill is plausible because Mill assumes that we value happiness as the ultimate end; Mill thought happiness was valuable because we all desire it (Mill 1879 49-50). It follows from this that if we were to all value something else rather than happiness, it would be that which would have ultimate value. Mill is not the only one to think this, and utilitarians generally think that utilitarianism applies to all sentient beings (Sidgwick 257). So, it seems that both Mill and utilitarians generally intend for utilitarianism to apply universally, and it would be a serious problem for utilitarianism, even amounting to a major internal inconsistency, if it did not.

In order to demonstrate the problem classical utilitarianism has with taking account of all free rational agents, let us engage in a thought experiment.

The thought experiment

Imagine a world of alien beings, who we will call neutralians, who possess free will, are self aware, have subjective conscious experiences, are capable of responding to reasons, but experience neither happiness nor unhappiness. Because they do not have the capacity to be either happy or unhappy, classical utilitarianism must consider any situation that they may get into morally neutral, unless it affects some other being that experiences happiness. According to a strict reading of classical utilitarianism, it seems that on the entire planet of Neutralia nothing of moral relevance ever happens. Mass murder would have the same moral status as a picnic in the park. If a human man were ever to go to Neutralia, then he would be the only being there with any moral status, because he would be the only one who could experience happiness or unhappiness. If the human wanted alien children to fight to the death for his amusement, then, so long as it made him happy, the neutralian children would be morally obligated to fight to the death. If the man began systematically exterminating neutralians, then, so long as it increases the human’s happiness and does not affect any other beings capable of experiencing happiness or unhappiness, classical utilitarianism would have to sanction this genocide.

These examples seem to show that utilitarianism cannot properly take account of all free rational agents. The fact that it grants some of them moral worth and others not, effectively denying morality exists on Neutralia, seems to pose a real problem for utilitarianism. Utilitarians can make several responses to this argument. Potential counter-arguments include: claiming that free rational agents must have the capacity to experience happiness; redefining what is meant by “happiness” in various ways; claiming that we cannot imagine a world such as Neutralia; and, of course, biting the bullet.

¹⁶ I will assume he meant having consciousness or subjective experiences, as this is what is generally meant by the word “sentience”

The utilitarian's responses

Disputing the premise

One response that utilitarians can make is to claim that the ability to experience happiness and unhappiness is necessary for creatures to be free, rational, self reflective agents. In order to do this, utilitarians could claim that it is an empirical fact that agents of this sort are capable of experiencing happiness or unhappiness.

Any argument that claims that it is an empirical fact that agents are capable of experiencing happiness or unhappiness misses the point of this argument. Because this is a thought experiment designed to show that utilitarianism cannot take account of all potential moral agents, the scenario only needs to be logically possible. Whether something is physically unlikely, or even physically impossible, has no bearing on whether it is logically possible. So long as no contradictions are included in the thought experiment, we must maintain that it is indeed logically possible. The utilitarian might claim that without happiness and unhappiness neutralians could never have evolved to be free rational agents. This is also an empirical claim rather than a philosophical one and is not relevant to the thought experiment, but it also does not seem that this claim about evolution is obviously true. It is not so outlandish to imagine that neutralians might have evolved a desire to reproduce and feed themselves without happiness and unhappiness acting as middlemen. Also, because neutralians have been invented for the purposes of a thought experiment, they need not have evolved at all. It seems reasonable to imagine that it would be possible for a god or some other intelligent beings to have created neutralians with appropriate desires, but without the ability to experience happiness or unhappiness. So, to show that neutralians are not free rational agents, classical utilitarians must find some contradiction between neutralians being free rational agents and them not experiencing happiness and unhappiness.

Utilitarians could claim that it is necessarily true that rational beings can experience happiness and unhappiness. One way to support an argument of this kind would be by equating happiness with desire satisfaction, as economists often do, and claiming that without desire and motivation we can have no rationality (Martin 2012 15). Utilitarians could argue that happiness and unhappiness are just terms for desired and undesired brain states, so we cannot desire anything, and therefore be motivated to do anything, without happiness and unhappiness. If the utilitarian defines "happiness" as desire satisfaction, then they can include happiness as an essential part of agency. This is because without the faculty for desire satisfaction, there would, presumably, be no desire or motivation. This matters because the kind of rationality we are interested in here consists of being able to respond appropriately to, and be motivated by, reasons. It seems that if one could not desire or be motivated, then one could not truly be rational in this way. Therefore, if one can indeed only have desires or be motivated if one can experience happiness, there would be a contradiction in the statement "a free rational agent incapable of experiencing happiness or unhappiness". However, equating happiness with desire satisfaction has significant flaws.

Classical utilitarianism justifies happiness as ultimately valuable by claiming that we all desire it as our ultimate end (Mill 1879 49-50). However, if happiness were equated with desire satisfaction, the utilitarian's claim cannot be false, because they

would be claiming that everyone wants happiness, but that happiness just amounts to desire satisfaction. This means that their claim would be that everyone wants their desires satisfied, which is like claiming that everyone wants what they want. The claim that everyone wants what they want is obviously true a priori; it is a tautology. This response also seems to reduce classical utilitarianism to a form of preference utilitarianism. If happiness and desire satisfaction were the same thing, then classical utilitarians would be preference satisfaction utilitarians under a different name.

Let me reiterate that I am not criticising preference utilitarianism in this chapter. The point I am making is that classical utilitarians cannot claim that happiness is the thing which has ultimate value, rather than preference satisfaction, and then claim that what they really meant by “happiness” was preference satisfaction when they run into problems. Since this thought experiment aims to show that the claim that happiness is the only non-instrumental good is inadequate, if classical utilitarians respond by altering their position on what has ultimate value, then the thought experiment has been successful.

Utilitarians may respond that they are not equating happiness and desire satisfaction, just claiming that they are inextricably linked. But, linked in what way? If they mean linked in some kind of a priori logically inseparable way, then it seems their claim is wrong. As previously shown, we do seem to be able to coherently imagine a type of being like a neutralian that has the faculty for desire satisfaction without happiness or unhappiness, and if we can coherently imagine it, it is logically possible. If, however, they mean that they just are linked as a matter of fact, then we are back to unsubstantiated empirical claims that are, as previously mentioned, irrelevant to the thought experiment.

Redefining happiness

Another response the utilitarian could make to the criticisms raised in this chapter is that happiness has been poorly defined in the neutralian thought experiment and, when happiness is properly defined, utilitarianism can take account of beings such as neutralians. In the neutralian thought experiment, it is assumed that happiness is a state of mind, or several states of mind, or something common to many states of mind, but not all definitions of happiness need refer to mental states. “Happiness” can also be defined as a life going well. This definition may be more likely to be universally applicable, but it is also vague. What qualifies as a life going well, and who decides whether a life is going well? Three of the most popular ways of defining a life going well are: theories based on pleasure; those based on desire; and objective list theories (Parfit 4). The suggestion that happiness can be defined in terms of desire satisfaction was addressed in the preceding section. So, rather than addressing it again, I examine only whether a life going well can be defined in terms of pleasure or using an objective list of criteria.

Pleasure

If we were to define a life going well in terms of pleasure, and define happiness in terms of a life going well, then we are defining happiness in terms of pleasure, or rather pleasure and the absence of pain. Many utilitarians use the word “happiness” instead of “pleasure” to avoid the simplistic connotations of the word pleasure, and defining “happiness” as meaning “pleasure” seems to be a step backwards. More importantly, this definitional manoeuvre does not help utilitarians out of the dilemma posed by the neutralian thought experiment, as it seems as easy to imagine beings without the capacity to experience pleasure, as it does to imagine beings without the capacity to experience happiness.

Objective list theories

Objective list theories of what it means to have a good life, and therefore what happiness is, are based on objective lists of criteria.¹⁷ The issue with objective list theories is what makes it onto the list. Lists may be species specific, such that what it means to lead a good life is different for a human than it is for a neutralian. Many people might agree that things like love and friendship are important to a life going well, but that may be simply because we are hyper-social animals and these things may not apply to neutralians, who may not be. However, even if there were a list for every different species, the list might not apply to every member of that species. Many people value doing good works with their lives, but some certainly do not. Most people value having a loving family, but there are those who do not. There are even those who do not value being in good health. Aside from perhaps vague things like “getting what you want”, which would take us right back to preference utilitarianism, it seems that whatever objective values you determine make a life good, there will be those who disagree. Even if we were to let every person have an individual objective list, it seems that there would still be people who got what is that they value, whether it is a successful career, a large family, or anything else, but who would still be miserable in the sense that most people would use the word (Martin 2008 176-180).

The problems I have raised for objective list theories of what it means to have a good life rely on identifying people for whom potential lists fail. The immediate response to this would be that, because it is an objective list, the things on it are things that would actually be good for someone rather than the things that they think would be good for them, but in actuality may not be. The problem with this response is how one defines “good”. Unless one defines “good” as “making someone happy”, there could easily be people who have an objectively good life but are still, by the layperson’s definition, miserable. It seems a problem for any definition of happiness if a person can be defined as happy and yet be in an emotional state that we would refer to as unhappy. This may lead some to claim that the things that are objectively good for people are the things that will make them happy, but this leads to a much bigger problem. If you were to define the objective goods in an objective list theory by reference to happiness, and define happiness as a life going well, and define a life going well by reference to an objective list of goods, then your argument would be hopelessly circular. It seems that utilitarians who are objective list theorists must either rely on a

¹⁷ Aristotelian virtue ethics is generally considered an objective list approach to wellbeing, and the objections here will apply to it, but Aristotle has rather more to say than just that wellbeing amounts to a list of objective goods, and will be discussed more fully in later paragraphs.

circular argument or define happiness in such a way that people can be objectively considered happy even though they are, from their perspective, miserable. Neither of these alternatives is appealing.

Let us take a look at the most commonly cited example of an objective list approach, the capability approach developed by Amartya Sen and Martha Nussbaum. The capability approach claims that what is important is that people have the freedom to pursue certain “functionings”, being and doing things that are of central importance to a human life (Alkire & Deneulin 32-37; Nussbaum 233). What this means is that capability theorists believe that the measure of a life going well is whether a person is substantively free to act in certain ways or make certain choices. Different philosophers espouse slightly different lists of specific capabilities that ought to be promoted but they tend to include things such as the ability to move around and to be well nourished (Sen 158). But, it is important to note that the capabilities approach does not claim that people ought to be free to do anything in particular, only that they ought to have several specific kinds of substantive freedom.

Leaving aside the problem of what should be on this list of capabilities, if utilitarians were to define happiness as a life going well and a life going well by reference to the capabilities approach, then their theory becomes a form of utilitarianism based on the promotion of a certain, kind of freedom. So, if the utilitarian tries to escape the problem posed in this chapter by redefining their theory using the capabilities approach, they are essentially saying that a form of utilitarianism which takes freedom to have ultimate value instead of happiness is better than one that uses happiness, which is the whole point of this thesis.

However, there is another way to define happiness, the Aristotelian way. When happiness is discussed in the Aristotelian tradition of ethics, it does not mean simply pleasure, it means living well and doing well (Aristotle 1095a), which can mean different things, including a life devoted to pleasure, a life devoted to honour and a life of reasoned contemplation (Aristotle 1176b-1179a). Aristotle thought that you could achieve great happiness even if your life contained no pleasure at all (1176b-1179a). If utilitarians were to adopt an Aristotelian definition of happiness, then they could say that the neutralians in this thought experiment are capable of experiencing happiness. Although the neutralians do not experience happiness in the sense that lay-people might use the word, they must be capable of experiencing at least some types of happiness, in an Aristotelian sense, in order to be free rational agents.

There are two major problems with using an Aristotelian definition of happiness as a response to the neutralian thought experiment, both of which I explain below. The first is that happiness defined in this way is clearly different to what classical utilitarians such as Mill thought of as happiness, and, secondly, were utilitarians to switch to an Aristotelian understanding of happiness, it would not work with their consequentialist framework.

First, let us analyse whether utilitarians actually do use “happiness” in an Aristotelian sense when they talk about happiness and its value. Some philosophers have argued that Mill uses some kind of Aristotelian account of happiness, and that his account of higher and lower pleasures makes most sense when thought of in this way (Kreider 53). However, when Mill talks of higher and lower forms of happiness, he talks of

higher and lower pleasures and the pleasure gained from various activities (Mill 1879 10-15). Claiming that some kinds of happiness are more desirable than others seems perfectly reasonable, without assuming he meant something entirely different from what he actually said:

Those who know anything about the matter are aware that every writer, from Epicurus to Bentham, who maintained the theory of utility, meant by it, not something contradistinguished from pleasure, but pleasure itself, together with the exemption of pain. (Mill 1879 9)

When Mill writes this, he really does seem to mean pleasure, in the sense that most people use the word.

If utilitarians did mean happiness in the Aristotelian sense, then they would have huge problems with their utilitarian calculations. Aristotle thought that happiness was the work of a lifetime and was not transitory in the way that we usually think of it (11). Under this definition of happiness, it would be very hard to weigh certain amounts of happiness against one another when making moral decisions. It is simply not a definition of happiness that lends itself to being measured and calculated. So, it would not be advantageous for the utilitarian to define happiness in an Aristotelian way because, while it would allow them to take account of free rational agents such as neutralians, it would not allow them to do the kind of utilitarian calculations for moral decision making that are the main strength of their theory. So, even if Mill did mean happiness in an Aristotelian way, and that is not the most common interpretation (Kreider 53), then he was wrong to do so, as an Aristotelian understanding of happiness does not fit with a utilitarian framework.

While it may seem initially attractive for the utilitarian to define happiness in an Aristotelian way, the theory would not be recognisably utilitarian. The point of this thought experiment is to bring to light a problem with classical utilitarianism, not a different, hybrid, theory, and it seems clear that when utilitarians talk about happiness they do not mean it in an Aristotelian way.

The biggest problem with any of these redefinitions of happiness is that they are defining “happiness” in a way that is different to how people, and utilitarians, actually use the term. People use and define happiness in many different ways, but one thing that is common to almost all definitions is that they refer to emotions or other brain states (Aaker, Kamvar & Mogilner 395; Haybron 501; Martin 2012 3; Mill 9). Defining happiness in a way that is completely divorced from not only Mill’s use of the term, but also the way in which it is used by both scholars and laypeople, makes the definition somewhat arbitrary. If utilitarians want to say that what they value is happiness, they must define “happiness” in a way that is connected to how the word is used and what word means. Otherwise they may as well say what they value is actually bananas, but that “bananas” is defined to mean a complex philosophical concept that can be changed ad hoc to solve any problems that arise.

The thought experiment is unimaginable

Another response the classical utilitarian can make to this chapter’s criticisms is to claim that the neutralian thought experiment posed is in some way unimaginable, or at least is not being properly imagined here. This is slightly different from claiming that there is some contradiction in the thought experiment, because it is not necessarily

claiming that the scenario in the thought experiment is impossible, it is just saying that we are not properly imagining it, so we cannot really debate it.

An example of this sort of objection is sometimes made to thought experiments involving philosophical zombies, where people are asked to imagine beings just like us only without consciousness (Marcus). The objector claims that the person or people involved in the imagining are not really imagining a lack of consciousness, they are simply imagining everything else and not imagining the consciousness. Because not imagining something is very different from imagining the lack of something, the thought experiment is flawed from the outset, as the subject of it is inconceivable. The main reason given for why we cannot imagine philosophical zombies is that we are being asked to engage in first person imagining, that is, imagining what it would be like to be something, and while we can imagine what it is like to be very cold or have a tail we cannot imagine what it would be like to have no consciousness. This is because our consciousness is the very thing doing the imagining. We may as well try to imagine what it is like not to exist (Marcus).

Claiming that the thought experiment is unimaginable may work for philosophical zombies but it does not for free rational agents who cannot experience happiness or unhappiness. While imagining what a lack of consciousness is like may be impossible, imagining a lack of happiness and unhappiness seems rather easy. In the case of zombies, you have never experienced what it would be like to not experience consciousness.¹⁸ In contrast, there are times where one is neither happy nor unhappy, and it is not such a great leap to imagine beings like neutralians who live their entire lives in such a state. The utilitarian may respond to this by saying that those times where you think you are neither unhappy nor happy are ones where you are actually just in a state of happiness or unhappiness so mild that you only think you experience neither. This seems a very odd claim to make though, as it seems that if I were experiencing happiness or unhappiness I would be in the best position to know it; claiming that I must happy or unhappy, even though I do not know it, seems like begging the question.

Biting the bullet

Another response the utilitarian can make to the claim that their theory cannot account for all potential free rational agents, is to bite the bullet. They can claim that agents like neutralians do not have moral status. They can claim that neutralians are free and rational agents, but that neutralians have no moral worth, because they experience no happiness or unhappiness. It should be mentioned that the utilitarian need not argue that the neutralian is not a moral agent. Utilitarians can say that neutralians are moral agents, because they are capable of making moral choices, but they have no moral worth because they cannot experience happiness or unhappiness. Utilitarians can argue that it is not a significant problem for their theory that it cannot take account of all free rational agents, and that taking account of all free rational agents is not the point of morality. Utilitarians can argue that they still have a good theory for regular human morality, or that our morality would not apply to aliens, or several other things that amount to saying that morality need not take account of all free rational agents.

¹⁸ Cases of driving home and arriving without seeming to have thought about it are better explained by memory and attentiveness issues than by a momentary lack of a conscious mind

However, it does seem that utilitarianism, as mentioned above, was intended to take account of all free rational agents, and perhaps a great deal more besides, and that its inability to do so is a problem for it.

Conclusion

In conclusion, it seems that classical utilitarianism cannot satisfactorily take account of all free rational agents. Free rational agency seems perfectly conceivable in beings that cannot experience happiness, but without the ability to experience happiness, classical utilitarianism has nothing to use as a measure of utility. The best solution to this seems to be to give up on happiness as the measure of utility and replace it with something else that is present in all potential free rational agents. In the final chapter I do just that, but in the next chapter I discuss how utilitarianism's focus on happiness, to the exclusion of all else, is both hypocritical, given the way in which utilitarians justify the value of happiness, and leads to a conflict with our personal integrity.

Chapter 4: Other values

Classical utilitarianism ignores many ways in which a life can be meaningful. It harms our integrity by insisting that when we pursue things such as knowledge or art at the expense of our happiness and the happiness of those around us, we are doing the wrong thing. I argue that given the method used by utilitarians to justify the non-instrumental value of happiness, ignoring these other things that people value for their own sake is hypocritical.¹⁹ In this chapter, I describe several situations where a person wants to sacrifice the happiness of themselves or those around them to pursue something else that they value. I look at how preference utilitarianism could deal with this and show that preference utilitarianism is also vulnerable to the objections raised in this chapter. Finally, I look at the other potential responses utilitarians can make and show that those responses cannot satisfactorily deal with the issue.

Demandingness

Before considering whether utilitarianism requires us to abandon our personal projects, we should first discuss the related objection that utilitarianism is too demanding. Many philosophers have criticised utilitarianism for demanding too much of us (Hills; Hurley; Mulgan). Utilitarianism is about maximising happiness, and, according to it, the right action to take is the action which produces the most happiness, or at least as much happiness as any other.²⁰ Critics take this to mean that, according to classical utilitarianism, it is wrong to do something that causes some happiness, but less happiness than some viable alternative action. We can use an example to demonstrate why this is a problem. Imagine you are faced with a choice between going to a movie and volunteering at the local school. Critics claim that in a situation like this, classical utilitarianism would say that neither of these is the right thing to do. The right thing to do is to sell any possessions you do not need to live or make money, donate that money to help those living in terrible poverty, and continue to give all but the bare minimum of your income to charity, while also participating in the most effective of the political and social movements designed to reduce global poverty. In fact, because the world is in a constant state of moral emergency where many people are in dire need of help and we are able to give that help, this seems, from a utilitarian perspective, like the right thing to do in almost every situation (Ashford 430).²¹ While everyone agrees that helping those in need to the extent that you can without damaging yourself is a very good thing to do, many people believe we should not be morally required to devote our entire lives to others. Morality should not require us all to be saints, this would be superogatory, that is, over and above the call of moral duty, rather than what we are all morally required to do at all times (Hills 225).

Utilitarians have several responses to the objection that utilitarianism is too demanding. For example, they can claim that overall it is better if everyone does good in their own community as their efforts will do the most good there. This response is

¹⁹ More precisely, I argue that it would be hypocritical given the method used by most utilitarians.

²⁰ Since utilitarianism is being treated as a form of expected value consequentialism in this work, it the expected consequences we are concerned with, rather than the actual ones.

²¹ Hills argues that because utilitarianism traditionally does not draw a distinction between happiness in persons and happiness in animals that are not persons, utilitarianism is actually even more demanding than this. However, the extent to which animals matter morally is not discussed here.

poor; it does not really address the demandingness objection at all, and it is demonstrably false given the extremity of the differences between rich and poor communities. The cost to feed someone living in poverty in most Western countries is much higher than it would be to feed someone living in poverty in Mozambique, so you get more good done for your dollar if you donate it to starving people in Africa instead of using it to help your local community.

Utilitarians could instead claim that although the best action to take may be that which maximises happiness, other actions may have good, but less good, consequences, and that one need not perform the best action to be morally praiseworthy, only a good one. This seems a good response as it coincides with the intuition many people have that to be worthy of praise one need not take the absolute best moral action, only a good one.

Philip Pettit responds to the demandingness objection in a similar way, saying that: "There is a distinction between what it is best to do and what you cannot reasonably be denounced for doing" (165). So, according to this response, while it may be best to give as much as you can to charity, you could not reasonably be denounced for not doing so. Pettit also claims that there is a difference between cases where you are the only person who could do something and cases where you are but one among many in the same position. Both of these claims seem intuitively sound. When it comes to assigning moral blame, many of us think it matters whether you are the only person who could take an action, or just one among many, and we do not denounce people for not always doing what would be best.

Although it makes some intuitive sense, the obvious objection to this response is that what you cannot reasonably be denounced for doing is not the same as what you ought to do. It could be the case that the reason we cannot reasonably denounce someone for failing to act in the best way is that we ourselves routinely act in ways that are far from the best, perhaps even ways that are actively bad, and it would be hypocritical of us to denounce someone for not being much better than us. If it is morally wrong not to take the best course of action, it does not become less wrong because everyone else is just as bad if not worse. This is the philosophical equivalent of arguing that you should not receive a speeding ticket because there were other people who were speeding even more than you.

Elizabeth Ashford presents another kind of response to the demandingness objection. Rather than denying that utilitarianism makes these kinds of demands of us, she readily admits it. However, she goes on to say that utilitarianism only makes such demands because the current state of the world is one of moral emergency, and that if we were to make large-scale political reforms, we could create a world where utilitarianism is not so demanding (Ashford 438). She also claims that compared to other theories, such as contractualism, utilitarianism is not particularly demanding. Ashford argues that, while utilitarianism could be made less demanding by improving the overall state of the world, contractualism, because it does not endorse certain kinds of trade-offs of welfare, would be extremely demanding in almost any practically possible state of the world (Ashford 298-301). This argument is different from the rest in that it does not seek to do away with the problem, but rather argues that there is no better, less demanding theory. However, the theory presented in this work, freedom consequentialism, is such a theory.

This chapter does not focus on the idea that utilitarianism demands too much of us in general, instead it focuses on the idea that utilitarianism demands that we give up on potentially valuable pursuits when they will not maximise happiness. The demandingness objection, and the responses to it, become relevant later in this chapter, but we will assume for now that utilitarianism can solve the demandingness objection using one of the responses listed above.

Examples

This chapter is primarily about classical utilitarianism failing to properly take account of things, besides happiness, that many people agree are worth pursuing for their own sake. While there are a number of candidates for things that people think are valuable, the ones focussed on here will be knowledge and art. These have been chosen primarily because they are things people value for their own sake, and they introduce fewer complicating moral considerations than values like justice or honesty.²² If we were to deal with honesty for example, then the issue at hand might be complicated by the considerations that go along with honesty. For example, whether being dishonest affects a person's future dispositions to do good things. So, to avoid overcomplicating the issue, the potentially valuable things chosen are knowledge and art.

Let us imagine a man. We will call him Simon. Simon is a scientist. He has been offered two different jobs, one doing research into the origins of the universe, and the other developing a way to remove the smell from farts. The research into the origins of the universe is only research into one highly technical aspect of the origins of the universe, and the general public will not be able to understand the findings of the research. So, assuming of course that the research into the origins of the universe does not yield any new technologies, it will likely yield a relatively small amount of net happiness. The research into removing the smell from farts, on the other hand, is likely to cause a large amount of net happiness, as people will no longer have to smell the farts of others or be embarrassed about the smell from their own farts. Let us also imagine that Simon is uniquely qualified to work on removing the smell from farts, and that if he does not take the job doing so, farts will continue to smell for many years to come.

Since one of Simon's options has the potential to bring about a lot more happiness than the other, classical utilitarianism would claim that it is the right choice and that taking the job researching the origins of the universe is wrong. However, many people would say it is a flaw in a theory if it maintains that it would be morally wrong for a scientist to study the origins of the universe instead of the smell of farts. The fact that classical utilitarianism does not recognise the pursuit of knowledge as a worthwhile way to spend one's life seems like a problem, because many people think that discovering the answers to life's most fundamental questions is valuable in and of itself. Whether there actually is any non-instrumental value to the pursuit of knowledge is up for debate, and not the point this chapter is trying to make. The relevant point is that what people value in their own lives matters.

²² I am not claiming that all people value these things for their own sake, only that some do. I also recognise that some may contend even this point. The response that we value these things for instrumental reasons is discussed later in the chapter.

This objection is very similar to the demandingness objection, as they both involve conflicts between making the most people happy and living the life one wants to. We have assumed that the utilitarian can escape that objection with one of the responses listed above. So, in the case of Simon the scientist, the utilitarian could respond with the same responses made to the demandingness objection. For example, they could claim that people should focus on doing what is best for their own community, and argue that those closest to Simon, such as the physics community and his loved ones, might be happier if Simon does something that he feels passionately about, such as taking the job researching the origins of the universe. However, it is not just the world at large that can be negatively affected by your choice to pursue things other than maximising happiness, it is also those closest to you. To illustrate this, let us engage in another thought experiment.

Let us imagine a man, we will call him Arty. Arty has a wife and many friends and relatives who care about him and enjoy his company. Arty has a good job that he enjoys and leads a happy life overall. However, Arty wants to be an artist. Being a great artist has been Arty's goal in life for as long as he can remember.²³ Arty is middle-aged, and he knows that if he does not leave his job immediately and devote his time to being an artist, then he will not have enough time to achieve his goal of becoming a great artist. He has a choice to make, he can give up on his dream or he can quit his job and pursue it. Arty's wife does not want him to quit his job. She wants to have children and wants Arty to have a well-paid job so that he can support his family. His friends and family look down on artists and would disapprove of his choice to become one. These factors, along with a desire to live somewhere inspiring, have lead Arty to decide that if he wants to become a great artist he would need to not only quit his job but also leave his family and friends behind and move away. It should be noted that if Arty finally gives up on his dream to become a great artist, he will be disappointed for a short time but he will get over it. Also, if Arty does leave his friends, family and wife to become an artist, then his life overall will be hard and lonely; he will gain some satisfaction from following his dreams, but he will not be as happy as if he had stayed. If Arty becomes an artist, his contribution to the art world will make the same number of people happy as those non-family members made happy by his current work. So, the choice Arty has to make is between a life where he and those closest to him are happy and one where he and those closest to him are unhappy but he pursues his dream of creating great art.

A classical utilitarian must say that it would be wrong for Arty to leave his family to become an artist. There are a number of potential reasons why this could be a problem for utilitarianism. One way to establish this is to argue that knowledge and art have non-instrumental value as well as happiness, and that ignoring them is a flaw in any moral theory. That is one potential way to look at this issue, but not the one examined here.²⁴ Instead, we will look at what Bernard Williams says on the matter. Williams argues that utilitarianism requires us to abandon our own projects in favour of pursuing the happiness of all, even our most fundamental projects that are at the heart of our identity, and he argues that this is highly damaging to our integrity (1993 108-

²³ This thought experiment is similar to the one used in the Gauguin argument (Williams 1993 37-38.) This thought experiment was used instead of the one in the Gauguin argument both to avoid arguments about what Gauguin's life was actually like, and so as to have a thought experiment tailor-made for examining utilitarianism's requirements rather than one designed to examine rational justifications.

²⁴ Developing a theory of aesthetic or epistemic value is beyond the scope of this work.

118). The utilitarian may respond that utilitarianism never set out to protect our integrity and that this is simply not a problem for it. However, as we have already shown, Mill's justification for happiness having moral value is based on the idea that we value happiness as our ultimate end (Mill 1879 49-50). This seems to indicate that it matters to utilitarians what our most fundamental projects or goals are, because their theory of what has value is based on what our ultimate end, or goal, is. So, if utilitarianism requires us to abandon those projects, goals or values that are most important to us, then this seems to be a flaw in the theory.

The utilitarians' responses

Classical utilitarians have several ways of responding to this issue. They can:

1. Bite the bullet.
2. Claim that the premise of the thought experiment is wrong.
3. Change their theory to preference utilitarianism, pluralist utilitarianism, or rule utilitarianism.
4. Claim that this objection is essentially just the demandingness objection and use the same responses used against that to solve this objection.

Let us examine these potential responses now.

Biting the bullet

One way for classical utilitarianism to respond to the criticism that it ignores other things that people consider valuable is to bite the bullet. The utilitarian can respond that utilitarianism does not give any weight to things such as art and knowledge, because they simply do not have any non-instrumental value and are only valuable to the extent that they bring about happiness.

There are several problems with biting the bullet. One is that it clashes with our intuitions. Many people have a strong intuition that pursuing knowledge or art for their own sake is a worthwhile endeavour.²⁵ Also, as discussed above, not allowing us to pursue our projects, especially those foundational to us as people, is highly damaging to our integrity (Williams 1993 108-118). Another problem with biting the bullet is that utilitarians claim, as mentioned earlier, that happiness is valuable because it is our ultimate end as people, so if some people have some other ultimate end, that other end would be valuable too. If Arty and Simon have artistic achievement and knowledge as their respective ultimate ends, then the utilitarian must concede that art and knowledge are to some extent valuable. In response to this, the utilitarian can claim either that people simply do not value art or knowledge as their ultimate end, or that people are simply wrong to value art or knowledge as their ultimate end. The claim that there are no people who value art and knowledge as their ultimate end seems a dubious one. If indeed there are people who are prioritising artistic achievement or the pursuit of knowledge over all else, including happiness, who is to say it is not their ultimate end? The utilitarian could claim they are mistaken about what their ultimate end in life is. Generally when we think people may be lying or confused about what it is that they really want, we look at their revealed preferences. In this case, if Arty were to leave his job to become an artist or Simon

²⁵ I am not endorsing this response, but it is worth mentioning that biting the bullet conflicts with this intuition.

were to go to work researching the origins of the universe, then it would seem that their revealed preferences were the same as their stated preferences.

There is another sense in which someone could have the wrong ultimate end. They could know what their ultimate end actually is, but have an ultimate end that it is wrong for them to have. Saying that people have the wrong ultimate end seems very odd.²⁶ There would need to be some criteria for what constitutes a correct ultimate end. We could perhaps create such a criterion based on what a rational person would want or what it is common to want, but these criteria also seem to fit with art or knowledge, as seemingly rational people value these things and it is not uncommon for people to value them. One could potentially base the criterion on what the majority of people have as their ultimate end, which we will assume for the sake of argument is happiness. However, the mere fact that most people have one thing as their ultimate end does not seem to be a convincing reason to believe that those who have a different ultimate end are in some way wrong. There is not a good reason to believe that people either could not have an ultimate end other than happiness, or that if they did, they should not. Because of the way in which classical utilitarians like Mill justify the moral value of happiness, it seems if people do indeed have ultimate ends other than happiness, then the utilitarian must concede that those ends matter.

The utilitarian could respond to this by saying that sometimes people desire or value things that are in fact not desirable or valuable, and that what is truly valuable is what we would value if we were fully rational and fully informed. One may desire to drink a certain liquid believing it to be water, but if the liquid is in fact poison, we could say that even though it is desired, it is not desirable. Brandt applies this kind of thinking to what people desire, saying that some desires are irrational and would not survive if the person in question were to know and firmly believe all the true propositions relevant to that desire (Brandt 1969 46). For example, we can imagine that someone may believe that serving God is non-instrumentally valuable, even to the extent that this is their ultimate end in life, but if they knew and firmly believed that God does not exist, then their ultimate end would change. The utilitarian can argue that, in cases such as this, what the person actually desires as their ultimate end does not have ultimate value, because a fully informed fully rational version of them would not desire it.

This response does not seem to preclude other potential ultimate ends as non-instrumentally valuable; it only argues that ultimate ends would not be non-instrumentally valuable if they were irrational. However, there seems no good reason to believe that valuing art or knowledge is irrational in this way. So, while the utilitarian may be able to argue that some people's ultimate ends are not ultimately valuable, because they are irrational, they do not seem to be able to make that argument in the case of art and knowledge. Utilitarians place value on happiness

²⁶ This could be read as not that odd at all; it is not odd to say something like "Do not pursue notoriety at all costs because it will not bring you happiness. It's the wrong ultimate end for you to have." Allow me to clear this up. If the person in question only pursues notoriety because they believe it will bring them happiness, notoriety is not their *ultimate* end. Their ultimate end is happiness. For something to be one's ultimate end, they must value it not only above all other ends, but also for its own sake, so telling someone they have the wrong ultimate end because it will not bring them happiness misses the point. The point is that they want their ultimate end *more* than happiness, and for its own sake, not simply because they think it will make them happy.

because people desire it as their ultimate end (Mill 1879 49-50). Without a good reason to show that people who value things such as art or knowledge over happiness as their ultimate end are being in, some way, irrational, it is hypocritical of them to treat these things as not ultimately valuable.

Disputing the premise

Utilitarians might try to refute my argument that their theory conflicts with our personal integrity by requiring us to give up our fundamental projects by objecting to the premise of the Simon and Arty thought experiments. They can claim that, as a matter of fact, people do value happiness as their only ultimate end, and that even when it seems they are pursuing knowledge or art or anything else instead of happiness, they are doing so for reasons that reduce to reasons about happiness and unhappiness. That is, they could claim that the people in these thought experiments are only pursuing art and knowledge because those things give them happiness or they believe that they will.

It is a very strong claim to say that everything people do is in the pursuit of happiness and one that is not supported by much evidence. It certainly seems like people do things in the pursuit of things other than the happiness of themselves and others. The utilitarian could say that even when it looks like someone is pursuing art or knowledge in favour of happiness, their reasons for doing so reduce to reasons about happiness and unhappiness. This would be a poor response, however, as it makes an unverifiable empirical claim about people. Also, the claim that when people seem to be doing things for reasons other than the pursuit of happiness, they are really doing them because they unconsciously believe it will lead to happiness is guilty of begging the question.

The utilitarian can also dispute the premise that the choice to become an artist or to research the origins of the universe would lead to less happiness overall. They could make the argument that although it is specified that it leads to less happiness in these thought experiments, in real life it would actually lead to more. In the case of Arty, they can claim that in the real world great art causes great happiness and that he would not need to move away from his friends, family and wife, as they would eventually come around if they truly cared for him. In the case of Simon, the utilitarian can claim that scientific discovery does cause significant happiness, not only in the scientific community, but also to those with an inquisitive disposition, and that a more complete understanding of our universe and its origins will likely lead to better technology for the public at large.

This response misses the point of the criticism levelled at utilitarianism. The situations in question could be tweaked to remove any objection that the utilitarian may raise as to whether pursuing the other potentially valuable thing will really produce less happiness. Even if things such as knowledge and art often do lead to happiness, it seems that there will be many times when they will not. There do seem to be legitimate clashes between happiness and other potentially valuable things, such as knowledge and art. For example, Carl Jung once said:

The artist's life cannot be otherwise than full of conflicts, for two forces are at war within him—on the one hand the common human longing for happiness,

satisfaction and security in life, and on the other a ruthless passion for creation which may go so far as to override every personal desire. (173)
Even if the conflict between happiness and art is not always this extreme, it certainly does exist.

This potential response by the utilitarian seems to be little more than another way to bite the bullet. Here the utilitarian is trying to say that knowledge, art and other things people value for their own sake are valuable because they lead to happiness. So, when the pursuit of these things does not lead to happiness, or leads to less happiness than the pursuit of happiness itself, utilitarianism will still have to judge the pursuit of these things as wrong. Since those are exactly the kinds of cases we are looking at here, it seems that this is not a response to this objection at all.

The utilitarian can also dispute the premise of these thought experiments by saying that it is unclear which of Arty or Simon's actions will really lead to the best happiness overall, and in cases where it is not clear, it is better to err on the side of autonomy. If this is the case, then there is not a conflict between utilitarianism and Arty or Simon's fundamental projects here. However, there are two major problems with this argument. One is that these thought experiments specifically state which choice will lead to more happiness. So, it is not inscrutable or unclear at all, it is very clear indeed. The utilitarian can respond by saying that in the real world cases are not so clear-cut and that often it is unclear whether one action over another will definitely lead to better consequences. This brings us to the second problem with this objection. If utilitarians cannot usually tell which action is at least most likely to be the best one to take, then their theory does not help us know what to do in moral situations. If the only times utilitarianism can be applied is in cases where it is absolutely clear that one option is terrible and the other is excellent, it is not a good moral theory. The utilitarian may respond that this is unfair, and that their theory can still be applied in many circumstances, just occasionally things are not clear. This is a fine response, but it brings us back to the point that in a significant portion of those circumstances utilitarians must say it is wrong for people to pursue their own projects instead of propagating happiness. The point is that the utilitarians cannot have it both ways. Either you cannot usually make judgments about which action it is better to take, in which case their theory is not a good one for determining how to act, or you can, in which case there is still the issue that utilitarianism requires us to abandon our most fundamental projects for the sake of happiness.

Alternative versions of utilitarianism

Mill's higher and lower pleasures

The utilitarian may respond to the criticism that it ignores other things that people consider valuable by saying that, although art and knowledge are only instrumentally valuable, some pleasures are more valuable than others.²⁷ Mill argues that some pleasures are higher than others and that everyone, or almost everyone, who is

²⁷ The idea that art and knowledge can be satisfactorily taken account of with an account of only instrumental value is raised and argued against earlier in this chapter. The issue of whether these things are instrumentally valuable or non-instrumentally valuable will not be contested again here. Instead, I will show that Mill's theory of higher and lower pleasure simply does not have anything to say about the thought experiments in this chapter.

competently acquainted with both will prefer the higher pleasures, even if that means experiencing a much smaller quantity of higher pleasures than could be provided by lower pleasures (Mill 1879 10-15). If one were to take this view, one could argue that the pleasure gained when creating art or uncovering the mysteries of the universe is qualitatively very valuable, and it would be consistent with Mill's utilitarianism to condone, or at least not condemn, Arty or Simon's decision to pursue their own projects.

Leaving aside the matter of whether Mill's theory of higher or lower pleasures is true, and, according to Ben Saunders this is still a much-debated topic, there are still several difficulties with using it as a response to this objection. One problem is that the cases presented here are not cases of one person choosing between higher and lower pleasures. It is not at all obvious that the pleasure associated with the creation of great art is higher than the pleasure associated with the bonds of family, friendship and love. There is no suggestion that those who have experienced both these pleasures prefer the pleasure of creating great art, and it seems intuitively more likely that the reverse is true. If Arty chooses to leave his family to pursue his dreams of becoming a great artist, he will not be choosing a higher pleasure for himself over a lower pleasure for himself and those he cares about. In fact, he may well be choosing a lower pleasure for himself over a higher pleasure for himself and those he cares about.

Another problem with introducing higher and lower pleasures theory to respond to the objections raised in this chapter arises from the difficulty of establishing comparative weights for higher pleasures against lower pleasures. Does a small amount of a higher pleasure always outweigh a large amount of a lower pleasure? Some philosophers have argued that Mill never meant for all higher pleasures to have lexical dominance over lower pleasures, and that it can sometimes be rational to choose a large amount of a lower pleasure over a small amount of a higher pleasure (Saunders). If this is the case, then it seems obvious that in the cases of Arty and Simon, utilitarianism must still recommend that they do what creates the greater happiness rather than pursuing their own projects; in both cases the quality of their higher pleasure is vastly outweighed by the sheer quantity of other people's, and potentially their own, lower pleasures.

Rule utilitarianism

A different way for the utilitarian to respond to the objection that their theory conflicts with personal integrity by forcing us to abandon our fundamental projects is to switch to rule utilitarianism. Rule utilitarianism is a form of utilitarianism that holds that the rightness of a moral action is determined, not by the consequences, or expected consequences, of the action itself, but rather by whether it conforms to a rule that is justified by its consequences (Emmons 226; Sobel 147). Rule utilitarianism is still a consequentialist theory, but the evaluation of consequences is moved from the consequences of individual actions to the consequences of following certain rules for acting. For example, a rule utilitarian faced with the choice of whether or not to kick a child would not need to do any utilitarian calculations about the situation. The rule

utilitarian could simply refer to the relevant rule for this situation, such as, “do not kick children”.²⁸

The rule utilitarian could easily claim that even though there may occasionally be conflicts between happiness and other things that are considered valuable, pursuing these things usually leads to a great deal of happiness. So, a rule utilitarian could say that because a rule saying, “pursue knowledge or artistic creation”, would lead to good utilitarian consequences, pursuing these things is the right thing to do.

The main problem with this response is in the framing of the rule in question. The rule could be something like, “pursue those things that, after careful consideration, you believe to be of value”. A rule of that nature would certainly connect with Mill’s intuition that we are the best judge of what is in our best interests, and that letting us pursue what we deem to be valuable is the best way of achieving happiness (Mill 1910 71). But, what about all the people who value things that are harmful to themselves and others. There are people who, after careful consideration, believe it would be a good thing to do things that harm many people. Do we really want to make a rule that would sanction sadists pursuing their goals of torturing children because they value the suffering of others? I suspect not. What if we were to make it a rule that people should pursue the things that they value so long as they do not cause net unhappiness to those people in their lives? This would simply bring us back to where we started, to the fact that often pursuing things such as art or knowledge can, as it does in the thought experiments provided in this chapter, cause unhappiness.

In order to escape the problems just described with defining the relevant rule, the rule utilitarian can make their rule more specific. However, as is always the problem with rule utilitarianism, if you make the rules too specific it collapses into act utilitarianism. David Lyons argues that rule utilitarianism always collapses into act utilitarianism, because for every rule imaginable there are cases where breaking that rule would bring about more utility, so a more sophisticated rule that incorporated these cases would be a better one. In order to account for all possible situations a person could be in, the rules utilised would have to be so complicated and specific that they amount to just being act utilitarianism.

Even if rule utilitarianism did not collapse into act utilitarianism, we have seen that we do not have reason to believe that it would include a rule or set of rules that allowed us to sacrifice the happiness of ourselves and others in order to pursue other things we value. So, rule utilitarianism is not a good response to this chapter’s objections.

Preference utilitarianism

Another way the classical utilitarian can respond to the objection that they ignore other things that people consider to be valuable, and that they require us to abandon our fundamental projects, is to switch to preference utilitarianism. Preference utilitarians claim that: “Whatever action satisfies more preferences, adjusted according to the strength of the preferences, that is the action I ought to take.” (Singer

²⁸ These rules can be constructed to allow for fringe cases, such as kicking a child to get it out of the way of a speeding bus.

1981 101). We can take this to mean that, according to preference utilitarianism, what is good is what satisfies the most, adjusted for strength, preferences, and what is bad is to cause preferences to be frustrated or left unsatisfied. This is the central tenet of preference utilitarianism, but, as we have already seen in the introduction of this work, there is more to defining a version of utilitarianism than that.

To see how preference utilitarianism works, let us borrow a thought experiment from Richard Hare's *Moral Thinking* (94-96). Hare asks us to imagine a man in a car who wants to park his car in a parking space. The problem is that there is already another person's bicycle in the parking space. The question at hand is whether the man with the car should move the bicycle in order to park his car in the space. Hare claims that we should treat this situation as though it were our own preferences that conflicted with one another. Imagine that it was your own bicycle that was in the way of your car and it was moderately inconvenient to move it, but highly inconvenient to not park your car there. In this case, when you weigh your preferences you find that you would much prefer to park your car there than not move your bicycle, so if you are rational, you ought to move your bicycle. Hare goes on to say that if you are fully aware of the other cyclist's preferences, you will acquire preferences of equal strength about what should be done to you if you were in his situation. The conflict is essentially between your own preferences, and from there you can resolve your preference conflict in the same way as you did when it was your own car and your own bicycle.

Not all preference utilitarians believe that you would actually acquire new preferences if you were fully aware of the cyclist's preferences, however, preference utilitarians do believe you can weigh other people's preferences each other and against your own (Singer 1983). Preference utilitarianism aggregates value in the same way that classical utilitarianism does. This means that it judges the total good overall to come out of an action, not the average per person. Preference utilitarianism is also agent-neutral, which means that, according to it, whether something is right or wrong is the same regardless of whose perspective it is looked at from. These are the important features of preference utilitarianism to note as they bear direct relevance to whether preference utilitarianism requires us to abandon our fundamental projects. We will treat preference utilitarianism as a form of direct consequentialism, as opposed to rule consequentialism, as well as a form of expected value consequentialism. These attributes are less relevant here but, for the sake of consistency, it is better if the kind of preference utilitarianism we examine is as close to classical utilitarianism as possible.

At first look, preference utilitarianism seems unaffected by my claim that classical utilitarianism requires us to abandon our fundamental projects and therefore conflicts with our person integrity. Preference utilitarianism values preference satisfaction instead of happiness. For this reason, it seems that choosing to pursue things other than happiness is not a problem for it. Intuitively it seems that if what the agent in question prefers is something other than happiness, then that is what they ought to pursue. So, pursuing things such as art or knowledge instead of happiness looks initially like it is consistent with preference utilitarianism

However, a related objection still applies to preference utilitarianism. Just as it is not only the individual's happiness we must weigh when deciding whether something is right by classical utilitarian standards, it is not only the individual's preferences we

must weigh when deciding whether something is right by preference utilitarian standards (Hare 109). So, if we were to apply preference utilitarianism to the case of Arty, we would take into account his preference to become a great artist, and weigh that against his friends, family, and wife's preferences for him to remain in their lives. So long as there are enough other people's preference to outweigh Arty's, it would be wrong for Arty to quit his job and move in order to pursue his goal of becoming a great artist. In the case of Simon the scientist, as long as more people would prefer that he worked on making farts not smell, it would be wrong for him to take the job researching the origins of the universe. Simply switching to preference utilitarianism does not resolve the conflict presented in these thought experiments. There is still a clash between the utilitarian greater good, and agents' desires to pursue the personal projects that are so important to their integrity.

Much as the classical utilitarian can respond to this chapter's objections with higher and lower pleasures, the preference utilitarian can respond with fully informed and rational selves. Brandt argues that a state of affairs is only good for an agent if "that person would want it if he were fully rational" (1969 268). This notion that someone's preferences are only worth satisfying if they would still hold them if they were fully rational allows the preference utilitarian to escape claims that it is good for a person to eat sand if they have a preference to do so. It can be taken even further by introducing the idea that what matters is what you would want if you were fully informed as well as fully rational. There are certainly many examples where it seems that our motivations change for the better when we become better informed. For example, I would no longer want to drink the glass of clear liquid in front of me if I learned that it was poison and not water (Loeb 1). This, and many other cases like it, have led some utilitarians to claim that an individual's good consists of what they would want if they were fully informed about their circumstances, as well as fully rational (Loeb). So, on this account, the preferences that matter are not those that people actually have, but those they would have if they were fully rational.²⁹

The preference utilitarian can apply this notion of fully informed selves to the objection here and claim that if all the people involved were fully informed and fully rational, some of their desires would be different. They could claim that the person in question, Simon or Arty, would want to take the course of action prescribed by preference utilitarianism if they were fully rational and fully informed. Alternatively they could claim that Arty's family and friends would understand his goals and support him if they were fully rational, and people would not care about whether farts smelled if they were fully informed and fully rational.

There are problems with both of these claims. To begin with, let us analyse the claim that Arty and Simon would no longer choose to pursue their goals if they were fully rational and fully informed. This claim assumes that fully rational, fully informed people would either no longer value art and knowledge to the same extent, or would no longer prioritise them over the satisfaction of the preferences of those closest to them. There is no compelling reason to believe that if people were fully informed and fully rational, they would place less value on art or knowledge. These things are valued by people who are very informed and rational; it does not make sense to

²⁹ Or what their fully informed, fully rational self would want for their actual, non-fully informed, non-fully rational self.

assume that a fully informed fully rational person would no longer value them, or value them as strongly.

However, Arty and Simon could still come to value the satisfaction of those around them more than they value art or knowledge without diminishing the value they place on art and knowledge respectively. They could simply come to value the preference satisfaction of others equally to their own. Arty, for example, could come to understand the emotional pain he will cause his loved ones if he leaves, decide that it would be selfish to pursue his dreams at the cost of his friends' and family's preferences, and thus change his mind. However, this claim seems to assume that Arty was in some way not aware of how much his leaving would hurt those he cares about, which seems unlikely. We, as humans, do not need to be fully informed or fully rational to understand the pain of loss or abandonment, or predict how leaving our friends, family, or spouse will affect them emotionally.

The preference utilitarian could also claim that if Arty were fully informed he would recognise that putting his own preferences ahead of the preferences of his family and friends is irrational, as he does not have a good reason to believe his preferences are more valuable than the preferences of others. If what Arty is doing is irrational, then he would not do it if he was fully rational. However, this amounts to utilitarians assuming the truth of their theory in mounting a defence of it, which is begging the question. While the preference utilitarian might argue that being fully rational involves treating like alike, and there is no compelling reason to value your own preferences above those of other people, this assumes that the person in question will not consider other things, such as the creation of great art, to have non-instrumental value. In Arty's case, he believes that creating great art has value for its own sake, and not simply because it satisfies his preferences. Even if Arty were fully informed and fully rational, the situation he is in would not simply be a matter of weighing his preferences against those of his family and friends. Because Arty values creating great art non-instrumentally, and we do not have a good reason to think he would no longer do so if he was fully informed and fully rational, it seems he can rationally choose to pursue the creation of great art even at the cost of the preference satisfaction of others. This response does not resolve the problem; there is still a conflict between utilitarianism, in this case preference utilitarianism, and the other things that people value non-instrumentally, in this case creating of great art.

The preference utilitarian could also claim that if everyone involved were fully informed and fully rational, their objections to Simon and Arty pursuing their goals would disappear. In the case of Arty, the preference utilitarian could argue that Arty's friends and family would support his decision to become an artist if they were fully informed and rational, and in the case of Simon that if people were fully informed and rational they would no longer care about something so trivial as whether their farts smelled or not. If being fully informed and fully rational would change the other people in these thought experiments' preferences, then there would not be a conflict, under a fully informed, fully rational self preference utilitarian model, between preference utilitarianism and Arty and Simon's fundamental desires. However, there is no good reason to believe that these other people would change their preferences if they were fully informed and fully rational. In the case of Simon, it seems perfectly rational to not want to experience or produce bad smells. In the case of Arty, it seems perfectly rational to prefer your husband, or friend, to remain with you and to

continue to lead the life that makes you happy. These are not irrational preferences, and there is no reason to believe that if people were fully rational and fully informed they would no longer hold these preferences, so this issue cannot be resolved by appealing to fully rational, fully informed selves.

Just as classical utilitarianism ignores other things that people consider to be valuable when they do not maximise happiness, preference utilitarianism ignores other things that people consider to be valuable when they do not lead to maximising preference satisfaction.

Pluralistic utilitarianism

The utilitarian can try to take account of other things that people believe have non-instrumental value by switching to some form of pluralistic utilitarianism. This could potentially resolve the conflict between utilitarianism and people's fundamental projects, as well as the apparent hypocrisy of assigning value to happiness but not to other things that people value non-instrumentally. George Edward Moore, for example, has a theory of utilitarianism that takes account of the values talked about in this chapter, though he calls them "beauty" and "truth" rather than "art" and "knowledge" (Moore 56). Pluralism about values can certainly be attractive, as it fits with how most of us actually think about moral decisions. Most people have moral intuitions that many different things are morally valuable, and that morality is not reducible to a single dimension such as the happiness-unhappiness dichotomy. However, there is a major problem with pluralism; pluralism, especially pluralistic utilitarianism, requires weighing different types of values against one another, and different kinds of values may well be incommensurable with one another. If different values were incommensurable with one another, it would not allow us to use them in the simple utilitarian calculations that are the core strength of utilitarianism.

The idea that values are incommensurable with one another is often argued for by pointing to the regret or loss an agent may rationally feel at having forgone the lesser of two options in favour of the greater (Klocksiem). It is argued that, although one option may be more valuable than the other, there is something in the lesser option that is not compensated or made up for in the greater option. Michael Stocker uses the example of attending a lecture in a crowded auditorium. Upon entering the auditorium you see that there are two empty seats for you to choose from, each next to one of your friends. It would be good for you to sit next to either friend, but one friend is the all things considered better option. Stocker argues that there may still be something good about sitting next to the lesser friend that is not made up for in sitting next to the greater friend, perhaps some special wit that the lesser friend has that is not possessed or made up for in the greater friend (Stocker 272–273). The idea is that if you can rationally feel regret over not choosing the lesser option over the greater one, then there must be something present in the lesser option that is not commensurable with what is present in the greater option. Stocker's example does not raise ethical issues, but it is easy to see how the same idea could be applied to moral decisions. So, although pluralistic utilitarianism can potentially solve the objections raised in this chapter, it creates much larger problems, such as not being able to do the utilitarian calculations that are one of the main strengths of utilitarianism because you are weighing incommensurable values.

Equating my objection to the demandingness objection

The utilitarian can also respond that the objection raised in this chapter, that utilitarianism requires us to abandon our fundamental projects and thus conflicts with person integrity, is really just another form of the demandingness objection, and can be solved in the same way. The problem with this claim is that it does not seem that the utilitarian's potential responses to the demandingness objection could be used as responses to this objection. Let us go through them one by one and see.

The first response to the demandingness objection we looked at was the idea that it is better to do good for your own community rather than for people thousands of miles away. Leaving aside the issue of whether this is a good response to the demandingness objection, it does not seem to be relevant here. While this response might arguably apply to the case of Simon the scientist, as most of the people whose farts his research will affect would not live in his community, it certainly does not apply to the case of Arty the aspiring artist. In the case of Arty, the people who are most affected are those closest to him, his friends, family and his wife, so claiming that he should do good in his own community rather than for people on the other side of the world does not do anything to resolve the conflict between utilitarianism and Arty's fundamental desires.

The next response we examined was to take an approach to utilitarianism where there are several good actions and, although one of them may be the best, doing any of them is still doing good. To illustrate this let us imagine a man who can afford to give at the most thirty dollars a week to charity. If we were to take this approach to utilitarianism it would still be best for him to do so, but if he were to give five dollars a week to charity instead, we could still say that he was doing good and not condemn him for doing the "wrong" thing. This response to the demandingness objection seems a sensible one, however it does not seem to apply to this objection. If Arty were to choose to leave his family, he would be choosing an action that will reduce the happiness of everyone involved over one that will keep the happiness of everyone involved at the same level, in other words, he will be actively doing something that utilitarians consider wrong. So even if we were to adopt this approach to utilitarianism Arty would still be considered doing something bad if he chose to pursue his dreams of becoming a great artist.

Phillip Pettit's response to the demandingness objection is that there is a difference between cases where you are the only person who can take an action, and cases where you are but one among many who could take the same action. This may be the case, but it does not seem to apply here. Only Arty can decide whether to keep his job and stay in the lives of his friends, family and wife, and Simon is uniquely qualified to remove the smell from farts. They are both in circumstances where they are the only ones who could take such an action, so Pettit's response does not apply.

So, we can see that although this objection is similar in some respects to the demandingness objection, it is not the same, and it cannot be solved in the same way. The responses the utilitarian can make to the demandingness objection do not help them avoid the conflict between utilitarianism and the other things that people value non-instrumentally.

Conclusion

Classical utilitarianism cannot satisfactorily account for other things, besides happiness, that seem to be valuable pursuits in life. It seems that we must accept that because of classical utilitarianism's commitment to happiness above all else, it will always ignore the value of lives spent pursuing other things that most of us would consider valuable, and this both brings it into conflict with our personal integrity and is hypocritical given the way in which utilitarians justify their claim that happiness has ultimate value. The responses that the classical utilitarian can make do not help resolve this conflict. Changing to preference utilitarianism might help avoid the hypocrisy of claiming that happiness has non-instrumental value because people value it as their ultimate end, but that other things, such as art and knowledge, that some people may value as their ultimate end, do not. However, it does not remove the conflict with the fundamental projects that are at the heart of our personal identity and integrity. Pluralistic utilitarianism is the only option that might resolve this conflict, but it comes with a number of problems of its own. Pluralistic utilitarianism may not allow for utilitarian calculations, which are one of the main strengths of utilitarianism, as different values may be incommensurable. In short, the utilitarian is faced with two significant problems and no good way of resolving both of them. However, in the next chapter I offer just such a way.

Chapter 5: Solutions

Freedom consequentialism, the theory outlined in the first chapter, can solve the problems presented in the last three chapters. In this chapter, I show that freedom consequentialism takes account of what is important about moral rights, applies to and takes account of all free rational agents, and allows people to pursue their own projects and preserve their own personal integrity, by changing the focus of utilitarianism from maximising happiness to protecting freedom. I also look at four potential objections to freedom consequentialism: that freedom consequentialism makes morality too cold; that this theory cannot be justified in the same well-established way that classical utilitarianism can; that preference utilitarianism does the same job; and the demandingness objection. I analyse these criticisms and show that they do not pose a significant threat to freedom consequentialism.

Moral rights

In chapter two, I argued that the claim that classical utilitarianism ignores moral rights is correct. This is a problem for classical utilitarianism, not simply because it clashes with a widespread intuition that moral rights matter, but also because of what gives moral rights their value in the first place. As discussed in chapter two, moral rights are considered to be valuable because they are the safeguards of our autonomy or freedom, which in turn is essential to our personhood. When rights theorists accuse utilitarians of reducing people to mere happiness vessels, their complaint is that utilitarians care only for the happiness, or unhappiness, that a person can feel, not for those things that make them a person in the first place (Frey 8; Williams 1973 45). So, the problem for classical utilitarians is that when they dismiss moral rights, or reduce them to mere hurdles to be overcome or tools for bringing about good utilitarian consequences, they are dismissing one of the things that make us moral agents in the first place, our power of self determination.

I argue that when rights theorists talk about what gives moral rights their value, they are talking about something very similar to the kind of freedom introduced in the first chapter of this work. I go on to argue that, although the theory presented here is a consequentialist one, it still preserves what is most important about moral rights.

Why rights matter

As mentioned in chapter two, rights theorists often justify the value of moral rights by reference to autonomy (Sumner 23). Some rights theorists define autonomy as only pertaining to the central pursuits of one's life, not the peripheral ones (Griffin 64). This means that someone can have a right to freedom of movement, but this right, and their autonomy, is not violated by forcing them to drive on a particular side of the road. However, using this view of autonomy conflicts with the idea that rights are absolute; that one either has a right to something or one does not, and that rights do not atrophy or dissipate from disuse or uninterest, (Griffin 61; Sumner 23). It might be argued that an action is not autonomous if the person performing it did not understand all the relevant facts before he or she acted. However, this approach to autonomy is hopelessly impractical, as it would require a kind of prescience that may well be impossible, and is certainly not possessed by ordinary people. These kinds of autonomy are not what moral rights protect. Rights protect the ability to perform

certain actions done from our most central desires, but also the ability to perform those actions on a whim, at random and without knowing all the relevant information. People do not lose the right to free speech because they have not carefully considered what they are going to say or because they do not have a strong interest in what they are saying. So, “autonomy”, as it is traditionally defined, is not what rights protect.

Other theorists justify the value of moral rights by reference to freedom, claiming that what gives moral rights their value is that they protect our freedom (Griffin 61). But, when rights theorists talk about freedom, they mean a limited kind of freedom. Under the modern conception of rights, moral rights protect an individual’s ability to do, or not do or have, or not have, certain things or make certain choices, but not others. Rights theorists agree that one has a right to live, or a right not to be killed, but that one does not have a right to have their name spelled out in twenty foot letters on the moon (Frey; Griffin; Osiatynski; Sumner; Raz 1985). The point being, that the freedom rights safeguard is not freedom to do anything, it is a limited kind of freedom.

The limited kind of freedom safeguarded by moral rights and the limited kind of freedom that freedom consequentialism protects are essentially the same thing. The details of my account and theirs differ, for example, most rights theorists do not talk about things that you already own. However, most of them imply something similar in their lists of what qualifies as a right.

Rights theorists may claim that there is a basic right to clean water, which may seem like a problem for the claim that freedom consequentialism can take account of moral rights, as not everyone owns clean water. However, when examined more closely, this right reduces to a right to be nourished and healthy if one chooses to be.³⁰ So, rights to clean water are contingent on our needing clean water to keep our body healthy. If humans suddenly developed the ability to sustain themselves without any need for food or water, or indeed any hunger or thirst, we would not claim that food and water were innate rights. So, people do not have a right to clean water, they have a right to be nourished, which fits with my account of freedom over one’s body. So, both freedom consequentialism and modern theories of rights are protecting the same thing, or something very similar, just under different names, and, when examined closely, what looks at first to be a conflict between freedom consequentialism and moral rights is not one.

What this means for moral rights

Freedom consequentialism is still a consequentialist theory, and it will still recommend the violation of individual rights in some cases. For example, it would still sanction killing one person to save many, because the freedom of one person is less important than the freedom of many. However, the things that is at the heart of moral rights, what gives them their value, that they safeguard freedom, is the same thing that is at the heart of freedom consequentialism. Freedom consequentialism does not view people as mere happiness vessels whose freedom does not matter, it places ultimate moral importance on protecting that freedom. Moreover, the capricious will

³⁰ It would, at least under the modern conception of rights outlined in chapter two, not be considered a violation of one’s rights to go on a hunger strike.

of the masses would never be sufficient to overrule the freedom of the individual, such as is the case in the Christians and the lions thought experiment, where the lives of a few Christians are weighed against the happiness, or desires, of a great many Romans. The only way this kind of violation would be justified is if not doing so led to an even greater violation of freedom, as is the case in trolley problems where the choice is between the death of one person and the deaths of five. While freedom consequentialism does not account for rights in the way that rights theorists would like, it, like rights based theories, is concerned with protecting the freedom that is central to our personhood.

Moral rights can be treated as freedoms to do, or not do, or have or not have, certain things, rather than safeguards of those freedoms. For example, claiming that a person has a right to life, may amount to saying that the person has a choice, that is, whether or not they want to live and continue to live, that belongs to them and them alone. If we were to think of rights in this way, then to violate someone's rights is to take away one, or more, of the choices that are rightfully theirs to make, or, to put it a different way, to violate their freedom over something that already belongs to them. Thinking of rights in this way, you could say of freedom consequentialism that it is a consequentialist theory of rights.

Free rational agents

In chapter three, I argued that classical utilitarianism cannot apply to and take account of all free rational agents. We saw that, in the case of neutralians, classical utilitarians had to either deny the premises of the thought experiment or concede that their theory would not consider these kinds of free rational agents to be morally valuable. In this section I argue that the best way for a moral theory to take account of, and apply to, all free rational agents is for it to be based on something that is common to all free rational agents, and that the best option for that is freedom.

Defining the terms

Before we discuss how moral theories can take account of all free rational agents, we need to define the term "free rational agent". This has already been defined in the introduction, but it is useful to discuss it again. The term "free rational agent" consists of three separate claims about the kind of thing we are talking about, let us go through them one by one now. When I write that an agent is "free" I mean that the agent has the ability to make choices that are not wholly determined by anything external to that agent and that the agent could have made a different choice in the same circumstances. "Rational", when used in this work, simply means being able to consider and respond appropriately to reasons. As for "agent", the word is used in this work to mean any self aware being, that is, a being that possesses consciousness and can reflect on that consciousness.

As mentioned in chapter one, for a moral theory to "take account of" all free rational agents it must treat them all as moral patients, that is, ascribe them moral value. For a theory to "apply to" all free rational agents, it must be possible for all free rational agents to engage with and follow the maxims of that moral theory; in other words, all free rational agents must be moral agents. So, for a theory to take account of and

apply to all free rational agents, it must treat them as both moral agents and moral patients.

Why being universal matters

We have already seen that classical utilitarianism is incapable of taking account of and applying to all free rational agents. This is because beings like neutralians are free rational agents even though they do not experience happiness or unhappiness, and if something cannot experience happiness or unhappiness, then classical utilitarians must ascribe no value to it, as happiness is the only thing they consider ultimately valuable. The only exception would be if neutralians were in some way instrumentally valuable, and if they were on a planet where they never interacted with beings that could experience happiness and unhappiness, they would not be. I have shown that this is a problem for classical utilitarianism because utilitarianism is intended to apply to at least all free rational agents, if not a great many other kinds of beings besides (Mill 1879 17).

Even if a classical utilitarian were to respond that classical utilitarianism is not intended to be universal, I believe it should be. There is a strong intuition that morality should be universal, not specific to one species or planet, and many philosophers have claimed that it is (Burnor & Raley 11; Kant 36-37; Mill 17). There are, however, a number of philosophers who do not agree (Capps; Lynch & Massey; Lee; Strong). It seems as though those who disagree that morality should be universal are talking about something different from those who believe that it should. Morality can be defined as, and is defined as for the purposes of this work, a code of conduct, or set of rules or maxims, by which all persons ought to abide. This kind of definition implies that any theory of morality ought to be universal. Opponents of the view that morality should apply universally generally think that the kind of morality defined here does not exist, and that morality means something quite different to this, such as a system of getting along with each other or a collection of cultural practices.

This work has indeed assumed that this kind of strong, objective, universal morality exists, and a powerful attack against freedom consequentialism may be to argue that it does not. However, the point of this work is not to defend moral realism about universal morality; this work argues that if morality exists and is universal, classical utilitarianism does a poor job of representing it, and proposes a theory that provides a better representation of universal morality.

Thought experiment

It seems that the best way for a moral theory to take account of all free rational agents equally is for your theory to be based on something that all free rational agents possess. Freedom consequentialism is just such a theory. Let us return to Neutralia and see using freedom consequentialism, rather than classical utilitarianism, would affect what we could say about the behaviour of neutralians.

Neutralia is a planet whose only inhabitants are a race of free rational agents that feel neither happiness nor unhappiness, known as the neutralians. We have already seen that classical utilitarians must dismiss these beings as not morally relevant and are

committed to saying that any act committed on Neutralia is morally neutral.³¹ On Neutralia, according to classical utilitarianism, theft, assault and even killing have less moral importance than a human drinking a glass of orange juice or taking a walk in the park. Now let us look at Neutralia again, this time using the protection of freedom as the measure of utility. Using freedom consequentialism, what could we say about the behaviour of a neutralian who, for example, stole from another neutralian?

Let us imagine such a neutralian now, we will call it Flarg. Flarg is a thief. One day Flarg steals a whatsit from another neutralian called Blurp. What can we say about this? Well, because we are using the protection of freedom instead of happiness as our measure of utility, the same things we could say about it if the same thing happened between two humans. Blurp has the freedom to do what it likes with its property, Flarg has violated this freedom by stealing Blurp's property and this is just as wrong on Neutralia as it is on Earth.³² If Flarg and Blurp were humans and Flarg had stolen the Earth equivalent of a whatsit, then that would have the exact same moral status as the theft would on Neutralia. It is not the unhappiness that a human Blurp would feel that determines the rightness or wrongness of the theft; it is the violation of Blurp's freedom over that which belongs to him. We can describe other actions in a similar way. The wrongness of murder is in the violation of the victim's freedom to live not in how many people grieve for them or how unhappy those people are, so it has the same moral status on Neutralia as it does on Earth.

What this means for free rational agents

Unlike classical utilitarianism, freedom consequentialism can apply to and take account of any logically possible free rational agents. This is because it measures utility by reference to the protection of freedom. "Freedom" is defined here as the ability to exercise one's free will, and free will must be present for something to be a free rational agent. Freedom consequentialism bases the measurement of utility on something that must exist in all free rational agents, so alien races, either real or imaginary, do not pose a problem for it.

Other values

As discussed in chapter four, classical utilitarianism considers only happiness to be non-instrumentally valuable, and ignores the other things that people value. This is a problem because, according to Mill, classical utilitarianism values happiness precisely because people value it as their ultimate end (1879 49-50), and if some people value something else as their ultimate end, then it seems inconsistent for utilitarians not to treat that thing as morally valuable as well. Also, in ignoring those things that people value besides happiness, classical utilitarianism requires people to abandon their own projects in favour of the happiness of others, even those projects that are foundational to their identity. Bernard Williams argues that that requiring us to give up these foundational projects is highly damaging to our personal integrity (1973 108-118). I

³¹ So long as no beings that can experience happiness or unhappiness are affected.

³² This assumes that Neutralians have a concept of property. If they do not, then there is obviously nothing wrong here. In societies that have no concept of property, we can think of everything they use as public property. So, while it would not be wrong for anyone to come and use it, it would be wrong for another group of people to come along and annex that property, thereby depriving that society of its use.

argue that both of these problems can be solved by using the protection of freedom as the measure of utility, instead of happiness.

Thought experiments

Let us revisit the cases of Arty and Simon, presented in chapter four, in order to demonstrate how freedom consequentialism differs from classical utilitarianism.

In chapter four we analysed the case of Arty, a man who wants to be an artist but knows that in order to do so he would need to leave his friends, family, wife and career behind and that this would cause a great deal of net unhappiness for him and those around him. Classical utilitarianism suggests that we ought to condemn this course of action, as it produces significantly more unhappiness than it does happiness. This assumes that art has no non-instrumental value; which it does not, according to classical utilitarians. According to utilitarianism, Arty is morally obligated to give up on his own projects to make other people happy. However, freedom consequentialism reaches a very different conclusion. In this case, no one's freedom is being violated. Arty's choice to either give up everything to become an artist, or give up on that dream and enjoy the life he has, is his and his alone. It is a choice over what to do with his life and therefore belongs to him. Because the freedom protected by freedom consequentialism is a limited kind of freedom, freedom over those things that belong to you, Arty can choose to pursue art without violating his friends, relatives and wife's freedom. Arty does not belong to them, and is under no obligation to continue his relationships with them. This does not mean it would be wrong for Arty to choose abandon his own projects to make those he loves happy, it simply means he is not morally obligated to do so. Whichever choice Arty makes, it is his choice to make, and morality, according to freedom consequentialism, has nothing to say on the matter. This assumes that Arty has no children. Were he to have children who depended on him for their food, health, education or wellbeing, then Arty might be obligated not to abandon his family.

We can also use freedom consequentialism to look at the case of Simon the scientist, discussed in chapter four. Simon wants to take a job researching a particularly obscure aspect of the origins of the universe, but he has also been offered a job working on removing the smell from farts. He is uniquely qualified to help remove the smell from farts; with him the project will likely succeed, without him the project will likely fail. If he takes the job researching the origins of the universe he will be an asset to the project but there are other scientists just as qualified who would be just as useful. Classical utilitarians must say that Simon is obligated to take the job trying to remove the smell from farts as this will bring a great number of people happiness, whereas the research into the origins of the universe will not bring about much happiness, and will get done with or without him. Freedom consequentialism reaches a different conclusion. Just as with Arty, the choice here belongs to Simon and Simon alone. Whether Simon wants to pursue research that he finds intellectually stimulating or research that will make many people happy is up to him, and morality has nothing to say on the matter.

What this means for other values

Freedom consequentialism would agree with classical utilitarianism that things like art and the pursuit of knowledge do not have non-instrumental moral value. It would instead classify them amongst the things that we value as humans, but are not universally morally valuable. This may seem as though it has the same problem as classical utilitarianism, but it does not. When classical utilitarianism denies that other things that we value are non-instrumentally morally valuable, it is being inconsistent, because it justifies the value of happiness by reference to what we value as our ultimate end, that is, for its own sake and most amongst all the things that one values. Because of this, when classical utilitarianism ignores other things that some people value over and above happiness, it uses special pleading. It says that happiness is valuable because it is our ultimate end, but, even though some people have other things as their ultimate ends, those things do not have the same value. Freedom consequentialism does not justify the value of freedom by reference to what people value, as their ultimate end or otherwise, but rather by reference to what is common to all free rational agents, and what is most important to the concept of morality and moral responsibility. So, it is not inconsistent for freedom consequentialism to say that things that people value for their own sake, such as art and knowledge, are not non-instrumentally morally valuable, because it does not claim that what people value determines what is morally valuable.

What this means for integrity

Bernard Williams believes that it is the consequentialist nature of utilitarianism that leads it to be so damaging to our personal integrity (1973 108-118). In *Utilitarianism for and Against* he describes situations where the consequences are significantly better if one does something that one is deeply opposed to, such as working on biological weapons or killing an innocent man, rather than if one holds to one's personal convictions (Williams 1973 41-42). He goes on to argue that when the utilitarian responds to this objection by saying we are being squeamish, and that is a problem with us rather than a problem with the theory, they are getting something fundamentally wrong. Williams argues that the problem with having to sacrifice our personal convictions for the utilitarian greater good is not that it is unpleasant to do so, it is:

Because our moral relation to the world is partly given by such feelings, and by a sense of what we can or cannot "live with," to come to regard those feelings from a purely utilitarian point of view, that is to say, as happenings outside one's moral self, is to lose a sense of one's moral identity; to lose, in the moral literal way, one's integrity. (Williams 1973 45)

The problem Williams is pointing at is no mere issue of having to do things one does not want to do, it is an issue of ignoring the importance of the things that make up the moral character of the agent in question. By treating one's personal convictions, or moral feelings, as objects of utilitarian value, we are ignoring what it is that makes that person a moral agent in the first place.³³

³³ This is a similar objection to the one that rights theorists often raise about utilitarianism reducing people to mere happiness vessels. Only, where rights theorists are concerned with a person's autonomy or freedom, Williams is concerned with their integrity.

The consequentialist nature of utilitarianism may be a contributing factor to the problem Williams is pointing out, but it is not the only one. Using happiness, or indeed preference satisfaction, as the measure of utility is another. Using one of these as the measure of utility means that whenever your own projects are set against the disapproval of others, so long as there are enough others, you are obligated to abandon those projects. This classical utilitarian approach also treats the moral feelings and personal convictions of people as merely another way to create happiness or unhappiness.

If we were to use the protection of freedom rather than happiness as the measure of utility, it would allow people much more freedom to pursue their own projects. Upsetting or disappointing those close to you would no longer affect what you were morally obliged to do. There would still be times when you would be obligated not to pursue your own projects, namely, when pursuing those projects would violate another person's freedom. However, I am sure even Bernard Williams would agree that it is worth damaging one's personal integrity by hindering their ability to pursue their own projects when that project is, for example, becoming an accomplished serial killer. Freedom consequentialism is not nearly so restrictive of one's ability to pursue one's own projects as classical utilitarianism, because, according to freedom consequentialism, it does not matter if pursuing those projects upsets or offends some people, only if it violates their freedom.

The obvious response to this is to say that freedom consequentialism ignores personal moral convictions just as much as utilitarianism, and that freedom consequentialism would also recommend doing things that may be in complete opposition to one's most fundamental moral feelings, such as killing the innocent man in Williams' example (1973 42). My response to this is twofold. First, while freedom consequentialism could require people to do things that are set against their most deeply held moral convictions, it would only require this in the direst of circumstances. It would not, for example, require the chemist in Williams' example to take the job researching biological weapons just to keep his family happy, but it would require him to take that job if his children are starving (Williams 1973 41-42).³⁴ While many of us may agree that our personal convictions are important, I suspect most of us would also agree that, when the situation is dire enough, our personal convictions ought to be sacrificed for the greater good. Second, freedom consequentialism does not ascribe value to personal moral convictions, but it treats the decisions they lead to, and the ability to make those decisions freely as all important. Although freedom consequentialism does not single out decisions made from fundamental moral convictions as more valuable than ones made from other motivations, it regards the ability to make one's own decisions, including to follow one's personal convictions or not, as the most important thing there is. Because of this, freedom consequentialism goes a long way to solving Williams' problem of personal integrity, if it does not solve it entirely.

³⁴ This assumes that other people could also develop these weapons and will take the job if he does not. If he was the only person with the proper skill set, or everyone else had already declined, then, depending on the likely consequences of developing these weapons, he may be obligated not to take the job even to provide for his family.

Objections

Classical utilitarians can respond to freedom consequentialism in a number of ways. They can claim that: freedom consequentialism makes morality too cold; that this theory cannot be justified in the same well-established way that classical utilitarianism can; that preference utilitarianism does the same job; and that freedom consequentialism is still vulnerable to the demandingness objection. I will now examine their possible responses and analyse whether or not they are problematic to freedom consequentialism.

This makes morality too cold

One of the most obvious objections to freedom consequentialism is that it makes morality too cold and impersonal. That it ignores the emotions that make us human in the first place, and that by concerning itself solely with the ability to make certain choices, it misses something important to morality. This is an appeal to intuition, specifically the intuition that our emotions matter morally.

My response to this is to bite the bullet. Emotions are very important to humans, but that is a separate realm of concern to what matters morally.³⁵ Every free rational agent should be able to engage with morality, regardless of their emotional range or empathetic ability³⁶. While emotions may be an important part of how we humans came to care about what it is to be moral, it is not necessarily the only way. Morality should be able to account for kinds of persons that are not as empathic or emotional as the average human. So, to the charge that this theory is the morality of psychopaths, I say yes, but that is a *good* thing.

Freedom is too hard to measure

One of the chief strengths of utilitarianism is its ability to reduce moral decisions to simple calculations of utility; by using the utilitarian calculus, complex moral situations can be, at least in theory, reduced to relatively simple equations of expected value.³⁷ Utilitarians might object to freedom consequentialism by saying that, unlike happiness, the protection of freedom either, cannot be quantified in the same way, or does not fit the current models for utilitarian calculus and would require an entirely new model.

I admit that the protection of freedom is more difficult to weigh up than happiness versus unhappiness, if you know what the likely amounts of happiness and unhappiness that will be produced by an action are. However, if what truly matters morally is people's freedom, then it does not matter if we have to rework our way of doing utilitarian calculations, we ought to change our theory to one that values freedom rather than happiness. Also, although it may take some work to adjust the current models of utilitarian calculus to use the protection of freedom instead of happiness, once done it would improve those models significantly. We would no

³⁵ Emotions may, of course, have instrumental moral value inasmuch as they affect our freedom.

³⁶ Because it does not require any emotional range or empathetic ability, Freedom consequentialism may be a useful tool to those exploring robot ethics.

³⁷ These will be calculations of expected value, as this work is treating utilitarianism as a form of expected value consequentialism.

longer have to rely on people reporting their own subjective experiences of happiness and unhappiness and then trying to quantify that as a numeric value, or, worse yet, simply assigning a numeric value to their subjective experience based on what we assume it is like. This is an advantage because we cannot know if the subjective experience the person is having is the same as what they are reporting. Someone may claim to be ludicrously happy at a certain moment but they may be lying, or they may be the kind of person prone to exaggeration, or they may be very happy relative to the rest of their lives but still not as happy as most other people because they are generally unhappy people. There is a similar problem with reporting whether one lives a generally happy life. Someone could report that they are happy in their life because they feel they ought to, perhaps because they have what they wanted most of their lives, be it a family, their dream job or anything else, or perhaps because their religious beliefs tell them that is how they ought to feel, yet be unhappy on a day to day basis. The short version of why it is an advantage to not rely on personal reporting of subjective experiences is that we do not have a good way of knowing what people are actually feeling, and people lie.

As to the objection that the protection of freedom cannot be quantified at all, this is false. In order to quantify the protection of freedom we would need to clearly define what kinds of freedom we possess. For example, under freedom consequentialism, we do not have the freedom to choose between tea or coffee if we do not own some tea and coffee. However, we do have the freedoms to do what we will with our property (including money) and to enter into contracts or agreements with mutually consenting persons. So, it would be a violation of these freedoms to outlaw the selling of coffee, but it would not be a violation of these freedoms if the world's coffee supply were to run low and you were not able to find any. Quantifying the protection of freedom would also require some sort of ordering of the importance of the different freedoms we possess, and that ordering might possibly be different from species to species, but there is no reason to believe that this is not possible.

So my response to the objection that the protection of freedom will be too hard to measure for utilitarian calculus is that even if it is hard, it is worth doing.

It cannot be justified in the same way

An attractive feature of utilitarianism is it has a well-established meta-ethical justification (Driver 59; Mandle 538). When one asks “why should I be good?” or “why should I follow this theory?”, the utilitarian can respond by saying that being good and following utilitarianism is just the rational expansion of your own ultimate end, and that it would be irrational for you to do anything else. The utilitarian can criticise freedom consequentialism by saying that it cannot be justified in the way that other utilitarian theories are, by reference to what people want as their ultimate end.

Many utilitarians justify what has ultimate moral value by reference to what everyone wants as their ultimate end. They claim that the only thing everyone wants for its own sake is happiness and the avoidance of unhappiness, and then go on to say that maximising happiness and minimising unhappiness is simply an extension of that, because one cannot make an exception of oneself (Mill 1879 49-50). Freedom consequentialism cannot use this line of reasoning to justify itself meta-ethically. When people ask “why be good?”, freedom consequentialism cannot respond that

doing so is simply the rational expansion of one's own most fundamental desires, as the protection of freedom is not most people's ultimate end. The utilitarian could argue that not being able to be justified in this way is a significant shortcoming of freedom consequentialism.

The problem with this objection is that the way utilitarians have of justifying their morality has a significant flaw. As discussed in chapter four, the Millian justification for utilitarianism ignores those people who value something other than happiness as their ultimate, or one of their ultimate, ends. There are people who report valuing things besides happiness for their own sake, such as knowledge and art, some even value these things more than happiness itself. One could say that they are mistaken about what they value, or that they are lying, and that they do in fact value happiness as their ultimate end. However, we do not have any reason to believe this except for an assumption that everyone values happiness as their ultimate end. So, claiming that these people are mistaken about what they value, or lying, would be a case of begging the question.

Happiness cannot be said to be everyone's ultimate end if there are people who value something else more than they do happiness. Happiness cannot be said to be the only thing which has ultimate value simply because everyone values it as their ultimate end, if not everyone does value it as their ultimate end. If there are some people, and indeed it seems that there are, who value other things more than, or even as much as, happiness, then the meta-ethical justification of utilitarianism is not a good one, and it is not a problem for freedom consequentialism that it cannot be justified in the same way.

Millgram's interpretation

Elijah Millgram has offered a different interpretation of Mill's argument. He claims that by using Mill's technique for correcting desires, we can say that if those who are experientially privileged desire something, then it is desirable for the average person, even if average people do not actually desire it themselves, because of their lack of experiential privilege (Millgram 298). To be experientially privileged in this area, people must have experienced an improvement to the general happiness at a cost to themselves and have also experienced a benefit to themselves at the cost of the general happiness (Millgram 299). Millgram goes on to say that we ought to count those who lived in the past and those who will live in the future in our determination of what the majority of the experientially privileged desire (300). He argues that because Mill thought that a society of natural utilitarians was inevitable, Mill thought that utilitarianism would be "proved true" in the future, as the majority of people who will ever live are likely to be yet to live, and he believed they would be utilitarians (Millgram 301). Millgram argues that this is perfectly consistent with Mill's views, as Mill thought that "*no* body of theory—logic and the mathematical sciences included—is a priori true or necessary", so it would not be so peculiar for Mill's views to be contingent in this way (Millgram 302).

The problems with Millgram's reading and interpretation of Mill here are significant. Before we discuss them, let us leave aside the issue of whether this is actually what Mill meant, because the issue at hand is whether utilitarianism can be justified in this way, not whether Mill did so adequately.

A significant problem with this reading of Mill's principle of utility is that it is contingent not just on what people actually desire, but on what other people, separated by miles and years, desire. Millgram's interpretation seems to suggest that it is not merely what I desire that is desirable for me, but what the majority of the experientially privileged desire. Even if I myself am experientially privileged and have come to my desires through careful reasoning and self reflection, they are, according to Millgram's reading of Mill, not desirable for me unless they are what the majority of experientially privileged agents' desire. This means I am in a position where I cannot know what is desirable for me, because the truth of whether something is desirable can only be known by considering the desires of everyone in an experientially privileged position throughout time.

This is a different claim from the one Mill makes when he talks of higher and lower pleasures. Mill says that people, if they could experience the, so called, higher pleasures with the proper knowledge or disposition, would agree that higher pleasures were indeed better than the lower pleasures. Whereas Millgram claims that what matters is the desires of *most* of the experientially privileged people who ever lived and will live. Millgram is essentially saying that even if it is the case that these experientially privileged people disagree with one another about what is desirable, one group is right by virtue of there simply being more of them. This deals not just with what people do desire, or with that they would desire given the opportunity or right frame of mind, but with what they *ought* to desire in some kind of metaphysically unclear way. Millgram does not offer any defence of why people ought to desire something just because the majority does.

In short, Millgram's reading of the principle of utility does not make it any more plausible.

Freedom consequentialism may not be able to use the same kind of meta-ethical justification as utilitarianism, as there is no suggestion that everyone desires the protection of their freedom as their ultimate end. However, that justification is a poor one to begin with, so this is not a significant shortcoming of freedom consequentialism. Also, even if Mill's or Millgram's meta-ethical justification of utilitarianism was persuasive, this would not mean their justifications were the only ones. The fact that freedom consequentialism cannot be justified in the same way does not mean that it cannot be justified at all.

Preference utilitarianism does the same job

Most prominent utilitarians now are preference utilitarians of some type (Brandt; Hare; Singer 1993), and most of the objections raised in this work are to classical utilitarianism. The utilitarian could claim that preference utilitarianism solves all the problems that this theory does, so this work is really just attacking an outdated form of utilitarianism instead of dealing with the more current one.

Preference utilitarianism has been introduced and explained in chapter four, so I will only briefly define it here. Singer sums up preference utilitarianism as "Whatever action satisfies more preferences, adjusted according to the strength of the preferences, that is the action I ought to take." (Singer 1981 101). So, for the purposes

of this work, preference utilitarianism will be defined as a form of utilitarianism that uses the satisfaction of the most preferences, adjusted for strength, as the measure for utility and the frustration, or dissatisfaction, of the most preferences, adjusted for strength, as the measure of disutility. Preference utilitarianism, when discussed here will be assumed to be a form of expected act consequentialism.

I reject the claim that preference utilitarianism solves all of the problems presented in this work. For example, the objection that utilitarianism forces us to abandon our projects, and thus damages our personal integrity, applies just as well to preference utilitarianism as it does to classical utilitarianism. Preference utilitarianism also does not solve the conflict between utilitarianism and moral rights.

Moral rights can be ignored just as easily for the sake of the preference satisfaction of the majority as they can be for the sake of the majority's happiness. The preference utilitarian can claim that their theory can take account of moral rights because people care about moral rights, so protecting them will be amongst their preferences. However, this means that moral rights are only protected so long as the majority think they should be. If the majority decide that something else, increased security for example, is more important than protecting moral rights, moral rights can easily be overridden. Preference utilitarianism can ignore moral rights, and because of this, the freedom that is at the heart of our personhood, just as much as classical utilitarianism does. Freedom consequentialism offers a potential solution to this conflict. The preference utilitarian could respond to this by saying that it is not a person's actual preferences that matter, but rather the preferences they would have if they were fully informed and fully rational (Loeb). On one account of the relationship between preference utilitarianism and rights, if people were fully informed and fully rational, conflict between our moral rights and our obligations as utilitarians would not arise. In cases of conflict, the person in question would understand what was morally required of them and waive their right. People would be willing to act in this way because if they were rational, they would not make an exception of themselves and would place the preferences of many above their own preferences.

There are several problems with this argument. The idea that it is not what people actually prefer that matters, but rather it is what some fully informed, fully rational version of them would prefer that has moral value, leaves us with a theory that is not very useful in the real world. It is very hard to determine what a fully informed, fully rational person would prefer, since there are no such people and never have been. Also, while it is all very well to say that people whose rights will be violated if agents act in accordance with preference utilitarian obligations would choose to waive those rights if they were only rational enough and informed enough, this does not change the fact that they did not waive those rights. The entire point of moral rights is that they protect the freedom of the individual to do, or not do, or have, or not have, certain things, even when that individual is doing something irrational or that most people would disagree with. To ignore this by claiming that violating people's rights is not an issue because they would agree to the violation if only they knew enough and were sufficiently rational, is to ignore the point of moral rights in the first place.

Preference utilitarianism offers no solution to the problems utilitarianism faces with personal integrity or with moral rights. Therefore, it is not an adequate alternative or response to freedom consequentialism.

The demandingness objection

While freedom consequentialism solves several problems with classical utilitarianism, it is still a form of consequentialism and still vulnerable to the demandingness objection. The “demandingness objection” is the objection that claims that utilitarianism demands too much of us; that is, because utilitarianism requires us to maximise utility, utilitarianism obligates us to do things that we would consider above and beyond the moral call of duty, actions that are morally praiseworthy, but not morally obligatory (Hills; Hurley; Mulgan). The demandingness objection could be applied to freedom consequentialism by arguing that people starving in Africa are having their freedom to live violated by our inaction, and that freedom consequentialism, like classical utilitarianism, requires you to give the maximum amount of money that you can to charity and live an impoverished life yourself. This is a problem as it means that freedom consequentialism would demand we all do things that we think of as being laudable but not obligatory.

This objection is a legitimate one. People have come up with a number of potential solutions for this problem, some of which I discussed in chapter four, but they are unsatisfactory. Although I believe the demandingness objection can be solved, changing the measure of utility from happiness to the protection of freedom alone does not solve it. However, the fact that this proposed change to utilitarianism does not solve all of its problems is not a reason not to accept it. If it solves just some of the problems with utilitarianism, and does not cause more of its own, then it is a good change to make.

Conclusion

Changing the measure of utility from happiness to the protection of freedom provides several advantages. Freedom consequentialism keeps what is most important about moral rights, that they safeguard our freedom, while still being a consequentialist theory. Freedom consequentialism applies universally, to all possible free rational agents, not just to human agents. Freedom consequentialism can say that things that people value, such as art, only have instrumental value without being hypocritical or severely damaging our personal integrity. These advantages add up to a significant reason to change the measure of utility from happiness to the protection of freedom.

Conclusion

Classical utilitarianism faces significant problems: it ignores moral rights; it cannot take account of all free rational agents; and its focus on happiness means that it dismisses the other things that people value for their own sake. These problems lead to conflicts with autonomy, personal integrity and inconsistencies with the way in which utilitarianism justifies the value of happiness. I have argued that these problems stem from using happiness as the measure of utility, and that they can be solved by instead using the protection of freedom as the measure of utility. Because it uses the protection of freedom as the measure of utility, freedom consequentialism can maintain what is most important about moral rights, namely, that they safeguard our freedom. It can also take account of all free rational agents, regardless of their emotional range, because it is based on something that is present in all free rational agents; freedom itself. Freedom consequentialism can also allow people to spend their lives pursuing the things that they value and still claim that things such as art and knowledge are not morally valuable without being internally inconsistent.

Freedom consequentialism is not without its problems, however. Two of the more significant problems are that it cannot be justified in same way as classical utilitarianism, and so needs a new way of justifying why people ought to follow it, and that it does not solve the demandingness objection. These problems are significant, but I do not believe them to be insurmountable. I have not provided a new theory of moral responsibility or a new meta-ethical justification for being moral. I have shown that there is good reason to believe that the protection of freedom is a better measure of utility than happiness or, indeed, preference satisfaction. Freedom consequentialism might be criticised on the grounds that the protection of freedom cannot be measured in the same way as happiness or preference satisfaction. However, I have argued that this is not the case.

Freedom consequentialism is in its infancy, and there is a lot of work left to be done. First, a method of measuring the value of different freedoms must be developed. Although determining how the value of different freedoms should be measured may be difficult, it is worthwhile; developing a system for measuring the value of different freedoms will lead to a better form of consequentialism, and one that does not rely on subjective reporting to determine how much utility is produced in a given situation. Second, a theory of moral responsibility that allows people to determine how much freedom one is obligated to protect must be developed. I have argued that this would also be a step towards solving the demandingness objection. There is still a lot of work left to do to make this a fully-fledged moral theory, but the significant advantages it has over classical utilitarianism, and even preference utilitarianism, are enough to make me think it is work worth doing.

References

- Aaker, Jennifer & Kamvar, Sepandar & Mogilner, Cassie. "The Shifting Meaning of Happiness" *Social Psychological and Personality Science* 2.4 (2010): 395-402.
- Alkire, Sabina & Deneulin, Séverine. "The Human Development and Capability Approach" in *An Introduction to the Human Development and Capability Approach*, ed. Deneulin, Séverine & Shahani, Lila. London: Earthscan 2009.
- Aristotle. *Nicomachean Ethics*. Translated by Sachs, Joe. Massachusetts: Focus Publishing, 2002.
- Ashford, Elizabeth 'Utilitarianism, Integrity and Impartiality', *Journal of Philosophy* 97 (2000): 421-439
- Baron, Jonathan. "Biting The Utilitarian Bullet" *Behavioral and Brain Sciences* 28.4 (2005): 545-546
- Brady, Michael S. "How to Understand Internalism". *The Philosophical Quarterly* 50.198 (2000): 91-97
- Brandt, Richard. "Rational Desires" *Proceedings and Addresses of the American Philosophical Association* 43 (1969): 43-64
- Brandt, Richard. "Utilitarianism and Moral Rights". *The Canadian Journal of Philosophy* 14.1 (1984): 1-19.
- Brandt, Richard. *A Theory of the Good and the Right*. New York: Prometheus Books, 1998.
- Broome, John. "Does Rationality Consist in Responding Correctly to Reasons?" *Journal of Moral Philosophy* 4 (2007): 349-374.
- Coeckelbergh, Mark. "Robot rights?: Towards a social-relational justification of moral consideration" *Ethics and Information Technology* 12 (2010): 209-221
- Corbett, Bob. *Moral Obligations to Distant Others*. At <http://www.webster.edu/~corbetre/philosophy/moral/others/distant.html>, 1995.
- Decew, Judith. "Brandt's New Defence of Rule Utilitarianism: Ideal Rules and The Motivation to be Moral" *PHILOSOPHICAL STUDIES* 43.1 (1983): 101-116.
- Driver, Julia. *Ethics: The Fundamentals*. Oxford: Blackwell Publishing, 2007.
- Ekstrom, Laura. "A Coherence Theory of Autonomy" *Philosophy and Phenomenological Research* 53 (1993): 599-616.
- Emmons, Donald C. "Act vs. Rule-Utilitarianism" *Mind* 82.326 (1973): 226-233.

- Frankfurt, Harry. *The Importance of What We Care About*. Cambridge: Cambridge University Press, 1988.
- Frey, R. *Utility and rights*. Oxford: Basil Blackwell, 1985.
- Griffin, James. *On Human Rights*. Oxford: Oxford University Press, 2008.
- Hampton, Jean E. "Naturalism and Moral Reasons" *The Authority of Reason*. Cambridge: Cambridge University Press, 1998.
- Hare, Richard. *Moral thinking: its levels, method, and point*. Oxford: Clarendon, 1981
- Haybron, Daniel. "Happiness and Pleasure" *Philosophy and Phenomenological Research* 62.3 (2001): 501-528.
- Hills, Alison. "Utilitarianism, Contractualism and Demandingness" *The Philosophical Quarterly* 60.239 (2009): 225-242.
- Holtug, Nils. "The Harm Principle" *Ethical Theory and Moral Practice* 5.4 (2002): 357-389.
- Hooker, Brad. *Ideal Code, Real World: A Rule-consequentialist Theory of Morality*. Oxford: Clarendon Press, 2000.
- Hume, David. *A Treatise of Human Nature*. New York: Dover Publications, Inc, 2003
- Hurley, Paul. *Beyond Consequentialism*. Oxford: Oxford University Press, 2009.
- Jung, Carl. *Modern Man in Search of a Soul*. Oxford: Routledge, 2001.
- Kagan, Shelly. *The Limits of Morality*. Oxford: Clarendon Press, 1989.
- Kant, Immanuel. *The Metaphysics of Morals*. Ed. Mary Gregor. Cambridge: Cambridge University Press, 1996.
- Kant, Immanuel. *The Groundwork of the Metaphysics of Morals*. Ed. Mary Gregor. Cambridge: Cambridge University press, 1997.
- Kant, Immanuel. *The Critique of Pure Reason*. Translated by Meiklejohn, J. Pennsylvania: Pennsylvania State University, 2010
- Klocksien, Justin. "Moorean Pluralism as a Solution to The Incommensurability Problem" *Philosophical studies* 153.3 (2011): 335-349.
- Kreider, S. "Mill on Happiness" *Philosophical Papers* 39.1 (2010): 53-68.
- Locke, John. *Second Treatise of Government*. Maestro Reprints, 2012.
- Loeb, Don. "Fully-information theories of individual good" *Social Theory and Practice*, 21 (1995): 1-30.

- Lyons, David. *Forms and Limits of Utilitarianism*. Oxford: Oxford University Press, 1965.
- Marcus, Eric. "Why Zombies Are Inconceivable" *Australian Journal of Philosophy* 82.3 (2004): 477-490.
- Martin, Mike. "Paradoxes of Happiness" *Journal of Happiness Studies* 9.2 (2008): 171-184.
- Martin, Mike. *Happiness and the Good Life*. Oxford: Oxford University Press, 2012.
- Mill, John Stuart. *Utilitarianism*. London: Longmans, Green, and co, 1879
- Mill, John Stuart. *On Liberty*. George Routledge & Sons limited, 1910.
- Morrisson, Iain. "Moral and Nonmoral Freedom in Kant" *The Southern Journal of Philosophy* 45.1 (2007): 129-148.
- Moore, G. E. *Principia Ethica*, Cambridge: Cambridge University Press, 1903.
- Mulgan, Tim. *The Demands of Consequentialism*. Oxford: Oxford University Press, 2005.
- Nussbaum, Martha. "Women and Equality: The Capabilities Approach" *International Labour Review* 138.3 (1999): 227-245.
- Nozick, Robert. *Anarchy, State, and Utopia*. Oxford: Basil Blackwell, 1974
- Osiatynski, Wiktor. *Human Rights and Their Limits*. Cambridge: Cambridge University Press, 2009.
- Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1986.
- Pereboom, Derk. "Kant on Transcendental Freedom" *Philosophy and Phenomenological Research* 73.3 (2006): 537-567.
- Pettit, P. *The Consequentialist Perspective in Three Methods of Ethics*. Blackwell: Oxford, 1997.
- Raz, Joseph. "Rights-Based Moralities" *Utility and rights*, ed. Raymond Frey. Oxford: Basil Blackwell, 1985.
- Raz, Joseph. "Explaining Normativity: On Rationality and The Justification of Reason" *Ratio* 12.4 (1999): 354-379.
- Sartre, Jean-Paul. *Existentialism Is a Humanism*. World Publishing Company, 1956.
- Saunders, Ben. "Reinterpreting the Qualitative Hedonism Advanced by J.S Mill" *The Journal of Value Inquiry* 45.2 (2011): 187-201.

Scheffler, Samuel. *A Rejection of Consequentialism*. Oxford: Oxford University Press, 1994.

Sen, Amartya. "Human Rights and Capabilities" *Journal of Human Development* 6.2 (2005): 151-166.

Sidgwick, Henry. "Utilitarianism" *Utilitas* 12.3 (2000): 253-258.

Silber, John. *Kant's Ethics*. New York: Walter De Gruyter, 2012.

Singer, Peter. *Practical Ethics: second edition*. Cambridge: Cambridge University press, 1993

Singer, Peter. *The Expanding Circle: Ethics and Sociobiology*. Oxford: Clarendon, 1981.

Sobel, Jordan Howard. "Rule-utilitarianism" *Australian Journal of Philosophy* 46.2 (1968): 145-165.

Stocker, Michael. *Plural and Conflicting Values*. Oxford: Oxford University Press, 1992.

Sumner, Leonard. "Rights Denaturalized" *Utility and rights*, ed. Raymond Frey. Oxford: Basil Blackwell, 1985.

Sunstein, Cass. "Moral Heuristics" *Behavioral and Brain Sciences* 28.4 (2005): 531-542.

Wellman, Carl. *An Approach to Rights: Studies in Philosophy of Law and Morals*. Dordrecht: Kluwer Academic Publishers, 1997.

Williams, Bernard. *Utilitarianism: for and against*. Cambridge: Cambridge University Press, 1973.

Williams, Bernard. *Moral Luck*. Daniel Statman (Ed.). Albany: State University of New York Press, 1993.